

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 13:00:37 ; Search time 34.5 Seconds
(without alignments)
11158.217 Million cell updates/sec

Title: US-10-029-345A-108_COPY_532_2532

Perfect score: 667
Sequence: 1 gttgcatcattgcccattgagat.....tggaaatcattgaggtctcc 2001

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Word size: 1

Total number of hits satisfying chosen parameters: 115352

Minimum DB seq length: 473

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xip
-O/cgrr 1/USPTO.spool_p/US10029345/runat_21062004_124616_5706/app_query.fasta_1.2183
-DB=PIR 78 -OPMT=fastran -SUPFIX=oligo.rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdt -LIST=75
-DOCALLGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=50 -MODE=LOCAL -OUTFMT=ptio
-NORM=ext -HEAPSIZE=500 -MINLEN=473 -MAXLEN=200000000
-USER=US10029345_@CGN 1.1.52 @runat_21062004_124616_5706 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCAPOP=6
-FCAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

PIR_78:*
1: pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.3	619	2 T15969	hypothetical prote
2	9	1.4	737	2 A34402	peptidyl-dipeptida
3	9	1.4	890	2 A97750	DNA mismatch repai
4	8	1.2	480	1 IBEG	hydroxymethylbilan
5	8	1.2	515	2 G86332	FeF9.1 protein - A
6	8	1.2	525	2 E96786	protein F10A5.13 l
7	8	1.2	526	1 P5XR10	outer capsid prote
8	8	1.2	545	2 P95325	probable arylsulfa
9	8	1.2	555	2 VGN211	cell fusion glycop
10	8	1.2	605	1 A35459	glucose oxidase (E
11	8	1.2	640	1 QOBEU2	U135 protein - hum
12	8	1.2	672	2 F84524	En/Spm-like transp
13	8	1.2	677	2 T22333	histidine ammonia-
14	8	1.2	692	2 T52120	acyl-CoA oxidase (

ALIGNMENTS

15	8	1.2	727	2 AB2353	hypothetical prote
16	8	1.2	752	2 H90135	cell division cycl
17	8	1.2	789	2 AE2688	Na+/H+ antiporter
18	8	1.2	789	2 H97469	probable NADH dehy
19	8	1.2	835	2 A27211	vira protein - Agr
20	8	1.2	1153	2 T21386	hypothetical prote
21	8	1.2	1286	2 T18734	hypothetical prote
22	8	1.2	1308	2 T15280	hypothetical prote
23	8	1.2	1716	2 T14103	probable DNA-direc
24	8	1.2	1717	2 T13961	DNA-directed RNA p
25	8	1.2	2512	1 MNWVS	nonstructural poly
26	8	1.2	2514	1 MNW82	nonstructural poly
27	8	1.2	2761	2 T29285	hypothetical prote
28	8	1.2	2943	1 RBH7AP	adenomatous polypo
29	8	1.2	2845	2 T49505	adenomatous polypo
30	8	1.2	473	2 B97073	beta-glucosidase l
31	7	1.1	474	2 D91075	6-phospho-beta-glu
32	7	1.1	474	2 C85920	6-phospho-beta-glu
33	7	1.1	474	2 H65051	6-phospho-beta-glu
34	7	1.1	475	2 S49886	probable membrane
35	7	1.0	475	2 B70143	fibronectin/fibrin
36	7	1.0	476	2 H89773	hypothetical prote
37	7	1.1	477	2 E97073	phospho-beta-gluco
38	7	1.1	477	2 G83048	probable two-compo
39	7	1.1	478	2 C86807	beta-glucosidase (
40	7	1.0	478	2 S17836	fatty-acyl-CoA red
41	7	1.1	479	2 F70573	hypothetical prote
42	7	1.0	479	2 T47415	transporter-like p
43	7	1.1	480	2 H86411	protein F1K23.12 l
44	7	1.1	481	2 H69593	6-phospho-beta-glu
45	7	1.1	482	2 AB1421	beta-glucosidase h
46	7	1.1	483	1 A31521	orphan receptor TR
47	7	1.0	483	2 F87700	rfa protein (limp
48	7	1.1	484	2 B33501	myosin heavy chain
49	7	1.1	486	2 AF1107	phospho-beta-gluco
50	7	1.1	486	2 A11468	phospho-beta-gluco
51	7	1.1	486	2 D89911	amino acid carrier
52	7	1.0	486	2 S67589	probable membrane
53	7	1.1	487	1 T32635	phosphoprotein pho
54	7	1.0	488	2 AE0001	conserved hypochet
55	7	1.0	489	1 VGXEMV	surface glycoprote
56	7	1.1	490	2 T02545	probable MYB fami
57	7	1.1	490	2 A24262	phase-1 i flagell
58	7	1.1	491	2 S49790	probable membrane
59	7	1.0	493	2 T43465	hypothetical prote
60	7	1.1	494	2 S16121	flagellin - Salmon
61	7	1.1	496	2 S55665	hypothetical prote
62	7	1.0	496	2 T19776	Na+/H+ antiporter
63	7	1.0	498	2 F89861	probable kinase pr
64	7	1.1	499	2 AF0073	hypothetical prote
65	7	1.1	501	2 F87550	hypothetical prote
66	7	1.0	502	2 H75290	hypothetical prote
67	7	1.1	504	2 F72744	protein F5011.6 l
68	7	1.1	505	2 B86258	probable aspartic
69	7	1.0	506	2 T07915	hypothetical prote
70	7	1.0	506	2 F86253	probable WD-40 rep
71	7	1.0	507	2 G84577	probable long chai
72	7	1.1	508	2 H95271	cytoxin (EC 3.4.2
73	7	1.0	509	2 S49349	conserved hypochet
74	7	1.0	509	2 D86911	glyceraldehyde 3-p
75	7	1.1	509	2 G82409	

RESULT 1
T15969
hypothetical protein F08B1.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #ext_change 20-Sep-1999
C/Accession: T15969
R/Chisoe, S.
submitted to the EMBL Data Library, July 1995

A>Description: The sequence of *C. elegans* cosmid F08B1.
A:Reference number: Z18439
A:Accession: U15969
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-619 <CH>
A:Cross-references: EMBL:U23178; NID:G726421; PID:G726422; PIDN:AA046719.1; CESP:F08B1.1
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F08B1.1
A:Introns: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3

Alignment Scores:
Pred. No.: 4.76 Length: 619
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x T15969 (1-619)

QY 733 CACGCTTACGTGGGATCTCCCGCTCC 759
Db 223 HiscYleuAlaGlyIleSerArgSer 231

RESULT 2
A34402
peptidyl-dipeptidase A (EC 3.4.15.1) precursor, testicular - rabbit
N:Alternate names: angiotensin I-converting enzyme; dipeptidyl carboxypeptidase I; peptidyl
C:Species: *Oryctolagus cuniculus* (domestic rabbit)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A34402; A60724; A36232; C18700
R:Kumar, R.S.; Kusari, J.; Roy, S.N.; Soffer, R.L.; Sen, G.C.
J. Biol. Chem. 264, 16754-16759, 1989
A:Title: Structure of testicular angiotensin-converting enzyme. A segmental mosaic isozy
A:Reference number: A34402; MUID:89380303; PMID:2550457
A:Accession: A34402
A:Molecule type: mRNA
A:Residues: 1-737 <KUM>
A:Cross-references: GB:J05041; NID:G164744; PIDN:AAA31155.1; PID:G164745
R:Sen, G.C.; Thekkumkara, T.U.; Kumar, R.S.
J. Cardiovasc. Pharmacol. 16(Suppl.4), S14-S18, 1990
A:Title: Angiotensin-converting enzyme: structural relationship of the testicular and the
A:Reference number: A60724; MUID:91155372; PMID:1705622
A:Accession: A60724
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 73-173 <SEN>
A>Note: Identical sequences were obtained for mRNAs from lung and testes
R:Chen, Y.N.P.; Riordan, J.F.
Biochemistry 29, 10493-10498, 1990
A:Title: Identification of essential tyrosine and lysine residues in angiotensin convert
A:Reference number: A36232; MUID:91104959; PMID:2176870
A:Accession: A36232
A:Molecule type: protein
A:Residues: 154-160;235-242 <CHE>
R:Iwata, K.; Lai, C.Y.; El-Dorry, H.A.; Soffer, R.L.
Biochem. Biophys. Res. Commun. 107, 1097-1103, 1982
A:Title: The NH₂- and COOH-terminal sequences of the angiotensin-converting enzyme isozym
A:Reference number: A90107; MUID:83048249; PMID:6291514
A:Accession: C18700
A:Molecule type: protein
A:Residues: 33-35, 'SN', '38-39', 'SS', 'FAEL', '737' <IMA>
A>Note: Several of the amino acids in reported are tentative
C:Comment: The pulmonary and testicular isoforms of this enzyme differ substantially in
C:Superfamily: mammalian peptidyl-dipeptidase A
C:Keywords: alternative splicing; peptidyl-dipeptide hydrolase; testis; transmembrane pro
Alignment Scores:
Pred. No.: 4.63 Length: 737
Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x A34402 (1-737)

QY 1334 CATAGCTTGTGTCCTCCATCCAGAGTA 1308
Db 22 HiscYleuAlaGlyIleSerArgSerVal 30

RESULT 3
A97750
DNA mismatch repair protein Muts [imported] - *Rickettsia conorii* (strain Malish 7)
C:Species: *Rickettsia conorii*
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C:Accession: A97750
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in *Rickettsia conorii* and *Rickettsia prowazekii*.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: A97750
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-890 <KUR>
A:Cross-references: GB:AE006914; PIDN:AAL02939.1; PID:G15619468; GSPDB:GN00173
C:Genetics:
A:Gene: mutS
C:Superfamily: DNA mismatch repair protein muts

Alignment Scores:
Pred. No.: 4.5 Length: 890
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x A97750 (1-890)

QY 1406 ATGCTGCTTCCTCCTTATCAGACTG 1380
Db 572 MetLeuAlaSerSerIleSerGlyLeu 580

RESULT 4
188G
hydroxymethylbilane synthase (EC 4.3.1.8) precursor - *Euglena gracilis*
N:Alternate names: porphobilinogen deaminase; pre-uroporphyrinogen synthase
C:Species: *Euglena gracilis*
C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
R:Sharif, A.L.; Smuth, A.G.; Abell, C.
Eur. J. Biochem. 184, 353-359, 1989
A:Title: Isolation and characterisation of a cDNA clone for a chlorophyll synthase enzy
nthesised with a very long transit peptide in *Euglena*.
A:Reference number: S06109; MUID:90005485; PMID:2477247
A:Accession: S06109
A:Molecule type: mRNA
A:Residues: 1-480 <SHA>
A:Cross-references: GB:X15743; NID:G18411; PIDN:CAA33759.1; PID:G18412
A>Note: part of this sequence, including the amino end of the mature protein, was confir
A>Note: 175-Ala was also found
C:Function:
A>Description: catalyzes the stepwise polymerization of four molecules of porphobilinoge
A:Pathway: porphyrin biosynthesis
A>Note: acting with uroporphyrinogen-III synthase (co-synthase), which cyclizes hydroxy
C:Superfamily: hydroxymethylbilane synthase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase; chlorophyll biosynthesis; chloroplast;
F:1-139/Domain: transit peptide (chloroplast) #status predicted <TMP>
F:140-460/Product: hydroxymethylbilane synthase #status experimental <MAT>
F:395/Modified site: dipyrrolyl-methanemethyl (Yys) (covalent) #status predicted

Alignment Scores:
Pred. No.: 4.63 Length: 737
Score: 9.00 Matches: 9

Pred. No.: 52.3 Length: 480
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x IBEG (1-480)

QY 1722 CTACGAGGACGACGACGCTTCTC 1745
DB 279 LeuArgArgGlnGlnSerLeuVal 286

RESULT 5

G86332
F699.1 protein - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #ext_change 09-Nov-2001
C/Accession: G86332

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Amen, N.E.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: G86332
A/Status: preliminary

A/Molecule type: DNA
A/Residues: 1-515 <STO>

A/Cross-references: GB:AE005172; NID:G10086502; PIDN:AA012562.1; GSPDB:GN00141
C/Genetics:

A/Map position: 1
C/Superfamily: Arabidopsis membrane-anchored cellulase KOR

Alignment Scores:

Pred. No.: 51.8 Length: 515
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x G86332 (1-515)

QY 374 GTGGCTTGTGAGTCTCTCGTT 397
DB 499 ValGlyLeuLeuSerLeuVal 506

RESULT 6

E96786
protein F10A5.13 (imported) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #ext_change 23-Mar-2001
C/Accession: E96786

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Amen, N.E.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: E96786
A/Status: preliminary
A/Molecule type: DNA

A/Residues: 1-525 <STO>
A/Cross-references: GB:AE005173; NID:G9369363; PIDN:AA07112.1; GSPDB:GN00141
C/Genetics:
A/Genes: F10A5.13
A/Map position: 1
C/Superfamily: Arabidopsis membrane-anchored cellulase KOR

Alignment Scores:

Pred. No.: 51.6 Length: 525
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x E96786 (1-525)

QY 374 GTGGCTTGTGAGTCTCTCGTT 397
DB 508 ValGlyLeuLeuSerLeuVal 515

RESULT 7

P5XR10
outer capsid protein VP5 - bluetongue virus (serotype 10, American isolate)
C/Species: bluetongue virus

C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #ext_change 16-Jun-2000
C/Accession: A25419; S10538

R/Purdy, M.A.; Ritzer, G.D.; Roy, P.
J. Gen. Virol. 67, 957-962, 1986

A/Title: Nucleotide sequence of cDNA clones encoding the outer capsid protein, VP5, of
A/Reference number: A25419; MUID:86198636; PMID:3009695

A/Accession: A25419
A/Molecule type: genomic RNA

A/Residues: 1-526 <PUR>
A/Cross-references: GB:D12532; GB:D01183; NID:G221078; PIDN:BAA02095.1; PID:G221079

R/Roy, P.; Marshall, J.J.A.; French, T.J.
Curr. Top. Microbiol. Immunol. 162, 43-87, 1990

A/Title: Structure of the bluetongue virus genome and its encoded proteins.
A/Reference number: S10534; MUID:90345726; PMID:2166648

A/Accession: S10538
A/Status: preliminary

A/Molecule type: genomic RNA
A/Residues: 1-526 <ROY>

C/Genetics:
A/Map position: segment 5
C/Superfamily: bluetongue virus outer capsid protein VP5

C/Keywords: capsid protein

Alignment Scores:

Pred. No.: 51.6 Length: 526
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x P5XR10 (1-526)

QY 98 CTATCATATGACGACATTTTCC 75
DB 435 LeuSerLeuSerLeuSerLeuVal 442

RESULT 8

F95325
probable arylsulphatase (imported) - Sinorhizobium meliloti (strain 1021) megaplasmid pS;
C/Species: Sinorhizobium meliloti

C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #ext_change 30-Sep-2001
C/Accession: F95325

R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Aboja, A.P.; Barloy-Hubler, F.; Bow
., Katman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A/Reference number: A95262; MUID:21396509; PMID:11481432

A/Accession: F95325
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-545 <KOR>
 A/Cross-references: GB:AE006469; PIDN:AAK65168.1; PID:g14523612; GSPDB:GNO0165
 A/Experimental source: strain 1021, megaplasmid pSYMA
 R/Galibert, F.; Finan, T.M.; Long, S.R.; Punler, A.; Aboja, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federpiet, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, M.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Weiler, D.H.; Wong, K.; Yeh, K.
 A/Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A/Reference number: A96039; PMID:21368234; PMID:11474104
 A/Contents: annotation
 C/Genetics:
 A/Gene: Sma0943
 A/Genome: plasmid

Alignment Scores:

Pred. No.:	51.3	Length:	545
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.24%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x F95325 (1-545)

QY 1593 CGTAGCTGCTGCTGCTGCTGCTGCA 1570
 |||||
 15 ArgGluValLeuLeuLagGly 22

RESULT 9
 VGNZ11
 cell fusion glycoprotein precursor - Sendai virus (strain C39)
 N:contams: fusion glycoprotein F1; fusion glycoprotein F2
 C/Species: Sendai virus
 C/Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jul-1999
 C/Accession: A31287; S11284
 R/Merson, J.R.; Hull, R.A.; Estes, M.K.; Kasel, J.A.
 Virology 167, 97-105, 1988
 A/Title: Molecular cloning and sequence determination of the fusion protein gene of human
 A/Reference number: A31287; PMID:89045674; PMID:2847427
 A/Accession: A31287
 A/Molecule type: mRNA
 A/Residues: 1-555 <MER>
 A/Cross-references: GB:M22347; NID:g332579; PIDN:AAA46800.1; PID:g332580
 A/Experimental source: strain C39
 R/Richardson, C.D.; Scheid, A.; Choplin, P.W.
 Virology 105, 205-222, 1980
 A/Title: Specific inhibition of paramyxovirus and myxovirus replication by oligopeptides
 A/Reference number: S11266; PMID:81016739; PMID:7414950
 A/Accession: S11284
 A/Molecule type: protein
 A/Residues: 113-132 <RIC>
 C/Genetics:
 A/Gene: F
 C/Superfamily: parainfluenza virus cell fusion protein
 C/Keywords: glycoprotein; membrane fusion; transmembrane protein
 F/1-21/Domain: signal sequence #status predicted <SIG>
 F/22-112/Product: cell fusion glycoprotein F2 #status predicted <FR2>
 F/113-555/Product: cell fusion glycoprotein F1 #status predicted <F1>
 F/113-138/Domain: transmembrane #status predicted <TM>
 F/498-519/Domain: transmembrane #status predicted <TM2>
 F/100,241,529,552/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:

Pred. No.:	51.2	Length:	555
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.24%	Indels:	0

DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x VGNZ11 (1-555)

QY 1193 ACCTATTGCTGCTTCCAGCCTG 1170
 |||||
 13 SerLeuLeuSerSerSerLeu 20

RESULT 10
 A35459
 glucose oxidase (EC 1.1.3.4) precursor [validated] - Aspergillus niger
 C/Species: Aspergillus niger
 C/Date: 31-Aug-1990 #sequence_revision 09-May-1997 #text_change 20-Oct-2000
 C/Accession: A35459; S05668; S14129
 R/Federick, K.R.; Tung, J.; Emerick, R.S.; Mastertz, F.R.; Chamberlain, S.H.; Vasavada, J. Biol. Chem. 265, 3793-3802, 1990
 A/Title: Glucose oxidase from *Aspergillus niger*. Cloning, gene sequence, secretion from
 A/Reference number: A35459; PMID:90154060; PMID:2406261
 A/Accession: A35459
 A/Molecule type: DNA
 A/Residues: 1-605 <FR>
 A/Cross-references: GB:J05242; NID:g166510; PIDN:AAA32695.1; PID:g166511
 R/Krichbaum, M.; Heilmann, H.J.; Wientjes, F.J.; Hahn, W.; Jany, K.D.; Gassen, H.G.; Sh
 FEBS Lett. 255, 63-66, 1989
 A/Title: Cloning and DNA sequence analysis of the glucose oxidase gene from *Aspergillus*
 A/Reference number: S05668; PMID:90005974; PMID:2792372
 A/Accession: S05668
 A/Molecule type: DNA
 A/Residues: 1-605 <KR1>
 A/Cross-references: EMBL:X16061; NID:g2356; PIDN:CAA34197.1; PID:g2357
 A/Note: part of this sequence, including the amino end of the mature protein, was confir
 R/Whittington, H.; Kerry-Williams, S.; Bidgood, K.; Dodsoworth, N.; Peberdy, J.; Dobson,
 Curr. Genet. 18, 531-536, 1990
 A/Title: Expression of the *Aspergillus niger* glucose oxidase gene in *A. niger*, *A. nidula*
 A/Reference number: S14129; PMID:91168301; PMID:2076553
 A/Accession: S14129
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-32 <WH1>
 A/Cross-references: GB:X56443; NID:g2354; PIDN:CAA39826.1; PID:g2355
 C/Genetics:
 A/Gene: gox
 C/Superfamily: alcohol oxidase
 C/Keywords: FAD; flavoprotein; oxidoreductase
 F/1-22/Domain: signal sequence #status predicted <SIG>
 F/23-605/Product: glucose oxidase #status experimental <MAT>
 F/43-72/Region: beta-alpha-beta FAD nucleotide-binding fold

Alignment Scores:

Pred. No.:	50.5	Length:	605
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	1	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x A35459 (1-605)

QY 386 AGTCTCTGCTGTTCCCTGACC 409
 |||||
 7 SerSerLeuValValSerLeuVal 14

RESULT 11
 Q08E02
 UL5 protein - human cytomegalovirus (strain AD169)
 C/Species: human cytomegalovirus, human herpesvirus 5
 A/Note: host Homo sapiens (man)
 C/Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
 C/Accession: S09798
 R/Chen, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Hornsneil, T.;
 M.; Barrell, B.G.
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
 A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus

A:Reference number: S09749; MWID:90269039; PMID:2161319
A:Accession: S09798
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-640 <CHE>
A:Cross-references: EMBL:X17403; NID:G59591; PIDN:CAA5394.1; PID:G59640
A:Note: possible protein-coding frames are given
A:Note: the DNA sequence was submitted to EMBL, December 1989, in computer-readable form
C:Superfamily: cytomagalovirus UL35 protein

Alignment Scores:
Pred. No.: 50 Length: 640
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x QGBE02 (1-640)

QY 1190 TTATTGCTGCTTCAGCCTGCT 1167

DB 547 LeuLeuLeuSerSerSerLeuSer 554

RESULT 12

F84524
En/Spm-like transposon protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: F84524

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentoc, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

Neues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MWID:20083487; PMID:10617197

A:Accession: F84524

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-672 <STO>

A:Cross-references: GB:AE002093; NID:G4115362; PIDN:AAD03364.1; GSPDB:GND0139

C:Genetics:

A:Gene: AT2G15070

A:Map position: 2

Alignment Scores:

Pred. No.: 49.7 Length: 672

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.24% Indels: 0

DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x F84524 (1-672)

QY 1094 ACCCTGGGCGCTGCGGATGC 1071

DB 347 ThrLeuGlyThrLeuAlaGlyCys 354

RESULT 13

T22333
histidine ammonia-lyase (EC 4.3.1.3) - Caenorhabditis elegans

N:Alternate names: histidase; protein F47B10.2

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T22333

R:Harris, B.

Submitted to the EMBL Data Library, November 1995

A:Reference number: Z19550

A:Accession: T22333

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-677 <WIL>

A:Cross-references: EMBL:Z68004; PIDN:CAA91982.1; GSPDB:GND0028; CESP:F47B10.2

A:Experimental source: clone F47B10

C:Genetics:

A:Gene: CESP:F47B10.2

A:Map position: X

A:Insertions: 24/3; 77/1; 111/2; 178/1; 255/1; 335/1; 399/1; 523/1; 544/3; 604/2

C:Superfamily: histidine ammonia-lyase

C:Keywords: ammonia-lyase; carbon-nitrogen lyase

F;269-271/Cross-link: 5-methylolone (Cys-Gly) #status predicted

F;270/Modified site: dehydroalanine (Ser) #status predicted

Alignment Scores:
Pred. No.: 49.6 Length: 677
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x T22333 (1-677)

QY 137 AAGCATTATATCAACTGCTCA 160

DB 584 LysProLeuIleSerThrAlaPro 591

RESULT 14

T52120
acyl-CoA oxidase (EC 1.3.3.6), peroxisomal [validated] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 03-Nov-2000

C:Accession: T52120

R:Hooks, M.A.; Kellas, F.; Graham, I.A.

Plant J. 20, 1-13, 1999

A:Title: Long-chain acyl-CoA oxidases of Arabidopsis.

A:Reference number: Z25966; MWID:20040045; PMID:10571860

A:Accession: T52120

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-692 <HOO>

A:Cross-references: EMBL:AF057043; PIDN:AACT13497.1

A:Experimental source: cultivar Columbia; seedling hypocotyl; 3 days old

C:Genetics:

A:Gene: ACX2

C:Function:

A:Description: EC 1.3.3.6 [validated, MWID:20040045]

A:Pathway: fatty acid beta-oxidation

A:Note: active with long-chain acyl-CoA; maximal activity with C18-CoA

C:Superfamily: acyl-CoA oxidase

C:Keywords: fatty acid beta-oxidation; oxidoreductase; peroxisome

Alignment Scores:
Pred. No.: 49.4 Length: 692
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x T52120 (1-692)

QY 596 AGCGTACCTTATCCCGAGTCTC 619

DB 34 SerLeuThrLeuSerProSerLeu 41

RESULT 15

AB2353
hypothetical protein al14378 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AB2353

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuch;

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, T.

DNA Res. 8, 205-213, 2001
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anaerostipes*
 A/Reference number: AB1807; MUID:21595285; PMID:11759840
 A/Accession: AB2353
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-727 <KUR>
 A/Cross-references: GB:BA000019; PIDN:BA876077.1; PID:G17133514; GSPDB:GN00179
 A/Experimental source: strain PCC 7120
 C/Genetics:
 A/Gene: all4378

Alignment Scores:

Pred. No.:	49.1	Length:	727
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AB2353 (1-727)

OY 1081 AGCGTGGCAGCGCGCCAGCGTG 1104

Db 116 SerValProSerValProSerVal 123

RESULT 16

H90135

cell division cycle protein 48 homolog [imported] - *Guillardia theta* nucleomorph

C/Species: nucleomorph *Guillardia theta*

A/Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 15-Jun-2001

C/Accession: H90135

R/Douglas, S.; Zauner, S.; Franholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reif, N.

A/Title: The highly reduced genome of an enslaved algal nucleus.

A/Reference number: A99082; MUID:11323671; PMID:11323671

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-752 <DOU>

A/Cross-references: GB:AF083031; NID:G13794396; PIDN:AK39773.1; GSPDB:GN00152

C/Genetics:

A/Gene: cdc48

A/Map position: 3

A/Genome: nucleomorph

C/Superfamily: transitional endoplasmic reticulum ATPase; Pfam/SECT8/CD48-type ATP-bind

C/Keywords: nucleomorph

Alignment Scores:

Pred. No.:	48.8	Length:	752
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.24%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x H90135 (1-752)

OY 771 AGCGATGCGGCGCGGAGAT 748

Db 570 SerAspGlyGlyAlaGlyasp 577

RESULT 17

AE2688

Nat/H+ antiporter mha [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)

C/Species: *Agrobacterium tumefaciens*

C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C/Accession: AE2688

R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, T.

erage, G.; Gillet, W.; Grant, C.; Guenhner, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan

; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001

A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

steir, E.W.

A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A/Reference number: AB2577; MUID:21608550; PMID:11743193

A/Accession: AE2688

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-789 <KUR>

A/Cross-references: GB:AE008688; PIDN:AAL41923.1; PID:G17739289; GSPDB:GN00186

A/Experimental source: strain C58 (Dupont)

C/Genetics:

A/Gene: mha

A/Map position: circular chromosome

Alignment Scores:

Pred. No.:	48.4	Length:	789
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.24%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AE2688 (1-789)

OY 1869 CTTTCAAAGGGGCTCTTCATG 1846

Db 340 LeuPheLysGlyAlaLeuPheMet 347

RESULT 18

H97469

probable NADH dehydrogenase (VCA0157) [imported] - *Agrobacterium tumefaciens* (strain C58)

C/Species: *Agrobacterium tumefaciens*

C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002

C/Accession: H97469

R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz, B.,

Science 294, 2323-2328, 2001

A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tun*

A/Reference number: A97359; MUID:21608551; PMID:11743194

A/Accession: H97469

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-789 <KUR>

A/Cross-references: GB:AE007869; PIDN:AKK6713.1; PID:G15155503; GSPDB:GN00169

C/Genetics:

A/Gene: AGR_C_1658

A/Map position: circular chromosome

Alignment Scores:

Pred. No.:	48.4	Length:	789
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.24%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x H97469 (1-789)

OY 1869 CTTTCAAAGGGGCTCTTCATG 1846

Db 340 LeuPheLysGlyAlaLeuPheMet 347

RESULT 19

A27211

vira protein - *Agrobacterium tumefaciens* plasmid pTiAg162

C/Species: *Agrobacterium tumefaciens*

C/Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 08-Oct-1999

C/Accession: A27211

R/Leroux, B.; Yanofsky, M.F.; Winans, S.C.; Ward, J.E.; Ziegler, S.F.; Neeter, E.W.

EMBO J. 6, 849-856, 1987

A/Title: Characterization of the vira locus of *Agrobacterium tumefaciens*: a transcriptio

A/Reference number: A91082; MUID:87246518; PMID:359559

A/Accession: A27211

A:Molecule type: DNA
A:Residues: 1-835 <IER>
A:Cross-references: GB:X05241; NID:g39191; PIDN:CAA28868.1; PID:g39192
C:Genetics:
A:Gene: vira
A:Genome: Plasmid
C:Keywords: autophosphorylation; host range; phosphoprotein; transcription regulation; v

Alignment Scores:

Pred. No.:	48	Length:	835
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.24%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x A27211 (1-835)

QY 545 ATCAGCTCTGTTGAGACATCT 522
|||||
DB 148 lleserSerleuValgthrSer 155

RESULT 20
T21386
hypothetical protein F29D11.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T21386; T21549
R:McMurray, A.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19415
A:Accession: T21386
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1153 <WIL>
A:Cross-references: EMBL:Z78419; PIDN:CAB01707.1; GSPDB:GN00019; CESP:F29D11.2
A:Experimental source: clone F26A3
R:Wilkinson, J.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19439
A:Accession: T21549
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1153 <W12>
A:Cross-references: EMBL:Z73907; PIDN:CAA98126.1; GSPDB:GN00019; CESP:F29D11.2
A:Experimental source: clone F29D11
C:Genetics:
A:Gene: CESP:F29D11.2
A:Map position: 1
A:Introns: 83/2; 117/3; 186/3; 243/3; 278/1; 353/1; 493/2; 543/2; 685/3; 884/2; 944/3; 1
C:Superfamily: Caenorhabditis elegans hypothetical protein F29D11.2

Alignment Scores:

Pred. No.:	45.7	Length:	1153
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.24%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x T21386 (1-1153)

QY 891 CTCATAGTCCAGAGTTGGCCG 868
|||||
DB 615 LeuileValGlnIleuAlaGln 622

RESULT 21
T18734
hypothetical protein B0391.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18734
R:Gardner, A.

submitted to the EMBL Data Library, November 1996
A:Reference number: Z19012
A:Accession: T18734
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1286 <WIL>
A:Cross-references: EMBL:Z81454; PIDN:CAB03805.1; GSPDB:GN00023; CESP:B0391.6
A:Experimental source: clone B0391
C:Genetics:
A:Gene: CESP:B0391.6
A:Map position: 5
A:Introns: 64/2; 277/3; 350/2; 427/2; 494/2; 522/2; 749/3; 806/1; 872/2; 898/2; 1111/3;

Alignment Scores:

Pred. No.:	44.9	Length:	1286
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x T18734 (1-1286)

QY 383 CTCAGTCTCTGCTGTTCCCTG 406
|||||
DB 1254 LeuSerSerleuValValSerleu 1261

RESULT 22
T15280
hypothetical protein R155.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15280
R:Geisel, C.; Wamsley, P.; Kramer, J.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid R155.
A:Reference number: Z18321
A:Accession: T15280
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1308 <GE1>
A:Cross-references: EMBL:AF003390; NID:G2088866; PID:G2088869; PIDN:AAB54272.1; GSPDB:G
A:Experimental source: strain Bristol N2; clone R155
C:Genetics:
A:Gene: CESP:R155.2
A:Map position: 3
A:Introns: 13/1; 57/1; 860/2; 897/1; 970/2; 1116/2; 1174/1; 1225/3

Alignment Scores:

Pred. No.:	44.8	Length:	1308
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x T15280 (1-1308)

QY 745 GGGATCTCCGCTCCGACCATC 768
|||||
DB 1165 GlyIleSerArgSerAlaThrIle 1172

RESULT 23
T14103
probable DNA-directed RNA polymerase (EC 2.7.7.6) I large chain - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
C:Accession: T14103
R:Hanman, R.D.; Hemple, W.; Rothblum, L.I.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z17875
A:Accession: T14103
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-1716 <HNV>
A:Cross-references: EMBL:AF025425; NID:g2739049; PID:g2739050; PIDN:AAB94601.1
A:Experimental source: Reuber hepatoma
A:Genetics:
A:Gene: RPA1
C:Superfamily: human DNA-directed RNA polymerase II largest chain
C:Keywords: nucleotidyltransferase

Alignment Scores:
Pred. No.: 42.9 Length: 1716
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x T13961 (1-1716)

QY 386 AGTTCCTCGTTGTTTCCTGGCC 409
|||||
Db 1537 SerSerLeuValValSerLeuAla 1544

RESULT 24

T13961
DNA-directed RNA polymerase (EC 2.7.7.6) I large chain - mouse
N:Contains: nonstructural protein NS1; nonstructural protein NS2; nonstructural protein
C:Species: Mus musculus (house mouse)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
C:Accession: T13961
R:Seither, P.; Coy, J.F.; Pouska, A.; Grummt, I.
Mol. Gen. Genet. 255, 180-186, 1997
A>Title: Molecular cloning and characterization of the cDNA encoding the largest subunit
A:Reference number: Z17838; MUID:97379976; PMID:9236775
A:Accession: T13961
A>Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-1717 <SE1>
A:Cross-references: EMBL:AF000938; NID:g2333006; PID:g2333007; PIDN:AAB6718.1
C:Genetics:
A:Gene: RPA1
C:Superfamily: human DNA-directed RNA polymerase II largest chain
C:Keywords: nucleotidyltransferase

Alignment Scores:
Pred. No.: 42.9 Length: 1717
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x T13961 (1-1717)

QY 386 AGTTCCTCGTTGTTTCCTGGCC 409
|||||
Db 1538 SerSerLeuValValSerLeuAla 1545

RESULT 25

nonstructural polypeptide - Sindbis virus
N:Contains: nonstructural NS2; nonstructural protein NS1; nonstructural protein NS3; non
C:Species: Sindbis virus
C>Date: 19-Feb-1984 #sequence_revision 03-Aug-1984 #text_change 04-Oct-1996
C:Accession: A03917
R:Straus, E.G.; Rice, C.M.; Straus, J.H.
Virology 133, 92-110, 1984
A>Title: Complete nucleotide sequence of the genomic RNA of Sindbis virus.
A:Reference number: A94331; MUID:8418439; PMID:6322438
A:Accession: A03917
A:Molecule type: genomic RNA
A:Residues: 1-2512 <STR>
A:Experimental source: strain HRSP

A>Note: readthrough of the terminator UGA between codons UAC for 1896-Tyr and CUA for 1
C:Superfamily: Semliki Forest virus nonstructural protein
C:Keywords: polypeptide
F:1-540/Product: nonstructural protein NS1 #status predicted <NS1>
F:541-1347/Product: nonstructural protein NS2 #status predicted <NS2>
F:1348-1896/Product: nonstructural protein NS3 #status predicted <NS3>
F:1897-2512/Product: nonstructural protein NS4 #status predicted <NS4>

Alignment Scores:
Pred. No.: 40.5 Length: 2512
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x MNV82 (1-2512)

QY 1550 CCAGCTTCCTTTGGCCTTCCA 1573
|||||
Db 445 ProLaserPheSerAlaPhePro 452

RESULT 26

MNV82
nonstructural polypeptide - Ockelbo virus (strain Eabym 82-5)
N:Contains: nonstructural protein NS1; nonstructural protein NS2; nonstructural protein
C:Species: Ockelbo virus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 08-Apr-1994
C:Accession: A39991
R:Shirako, Y.; Niklasson, B.; Dalrymple, J.M.; Straus, E.G.; Straus, J.H.
Virology 182, 753-764, 1991
A>Title: Structure of the Ockelbo virus genome and its relationship to other Sindbis vi
A:Reference number: A39991; MUID:91220725; PMID:1673813
A:Accession: A39991
A:Molecule type: genomic RNA
A:Residues: 1-2514 <SH1>
A:Cross-references: GB:M69205
C:Superfamily: Semliki Forest virus nonstructural protein
C:Keywords: nonstructural protein, polypeptide
F:1-540/Product: nonstructural protein NS1 #status predicted <NS1>
F:541-1347/Product: nonstructural protein NS2 #status predicted <NS2>
F:1348-1898/Product: nonstructural protein NS3 #status predicted <NS3>
F:1899-2514/Product: nonstructural protein NS4 #status predicted <NS4>

Alignment Scores:
Pred. No.: 40.5 Length: 2514
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x MNV82 (1-2514)

QY 1550 CCAGCTTCCTTTGGCCTTCCA 1573
|||||
Db 445 ProLaserPheSerAlaPhePro 452

RESULT 27

T29285
hypothetical protein C34D4.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29285
R:Du, Z.; Le, T.T.
submitted to the EMBL Data Library, May 1996
A:Description: The sequence of C. elegans cosmid C34D4.
A:Reference number: Z20600
A:Accession: T29285
A>Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-2761 <DUZ>
A:Cross-references: EMBL:U58755; PIDN:AAB00699.1; GSPDB:GN00022; CESP:C34D4.14

A:Experimental source: strain Bristol N2; clone C34D4
C:Genetics:
A:Gene: CBSP:C34D4.14
A:Map position: 4
A:Insertions: 120/1; 201/3; 365/3; 445/2; 509/1; 728/3; 856/3; 902/3; 940/3; 1027/3; 1508/3
Alignment Scores:
Pred. No.: 39.9 Length: 2761
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 2 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x T29285 (1-2761)
QY 1743 GTAACGGACAGCTGCTCGTAGAT 1720
Db 841 ValThrGlyThrAlaSerValAsp 848
RESULT 28
RBHUAP
adenomatous polyposis coli protein - human
N:Alternate names: polyposis coli locus protein DP2.5; tumor suppressor APC
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 21-Jul-2000
A:Accession: AJ37261; B39658; A44928; A49319; I54271
R:Kinzler, K.W.; Nibert, W.C.; Su, L.K.; Vogelstein, B.; Bryan, T.M.; Levy, D.B.; Smith
chul, S.F.; Horii, A.; Ando, H.; Miyoshi, Y.; Miki, Y.; Nishisho, I.; Nakamura, Y.
Science 253, 661-665, 1991
A:Title: Identification of PAP locus genes from chromosome 5q21.
A:Reference number: AJ37261; MUID:91335210; PMID:1651562
A:Accession: AJ37261
A:Molecule type: mRNA
A:Residues: 1-2843 <RIN>
A:Cross-references: GB:M74088; NID:G182396; PIDN:AAA03586.1; PID:G182397
R:Joshi, G.; Carlson, M.; Thilveris, A.; Albertsen, H.; Gelbert, L.; Samowitz, W.; Grod
arrington, J.; Mcherson, J.; Wasmuth, J.; Le Paslier, D.; Abderrahim, H.; Cohen, D.; Le
Cell 66, 601-613, 1991
A:Title: Identification of deletion mutations and three new genes at the familial polyp
A:Reference number: A39658; MUID:91330307; PMID:1678319
A:Accession: B39658
A:Molecule type: DNA
A:Residues: 1-183; 'L', 185-969, 'N', 971-11308, 'G', 1310-1324, 'SS', 1326, 'HSTLE', 1332-1354, 'P'
A:Cross-references: GB:M73548; NID:G190163; PIDN:AAA60354.1; PID:G190164
R:Miki, Y.; Nishisho, I.; Horii, A.; Miyoshi, Y.; Utsunomiya, J.; Kinzler, K.W.; Vogelst
Cancer Res. 52, 643-645, 1992
A:Title: Disruption of the APC gene by a retrotransposal insertion of Ll sequence in a c
A:Reference number: A44928; MUID:92119623; PMID:1310068
A:Accession: A44928
A:Molecule type: DNA
A:Residues: 1506-1525 <MIX>
A:Cross-references: GB:578214; NID:G243541; PIDN:AA21145.1; PID:G243542
R:Sprio, L.; Olschewang, S.; Groden, J.; Robertson, M.; Samowitz, W.; Joslyn, G.; Gelber
Cell 75, 951-957, 1993
A:Title: Alleles of the APC gene: an attenuated form of familial polyposis.
A:Reference number: A49319; MUID:9407973; PMID:8252630
A:Accession: A49319
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 'G', 143-171, 'P', 173-179 <SP1>
A:Cross-references: GB:567787; NID:G461061; PIDN:AA13997.1; PID:G426197
R:Lambert, S.; Ballhausen, W.G.
Hum. Genet. 90, 650-652, 1993
A:Title: Identification of an alternative 5' untranslated region of the adenomatous poly
A:Accession: I54271; MUID:9186137; PMID:8383094
A:Reference number: I54271; MUID:9186137; PMID:8383094
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-4 <LMB>
A:Cross-references: GB:556365; NID:G266243; PIDN:AA1918.1; PID:G4262770
C:Genetics:

A:Gene: GDB:APC
A:Cross-references: GDB:119682; OMIM:175100
A:Map position: 5q21-5q22
A:Note: mutations of this gene can result in familial adenomatous polyposis or sporadic
C:Superfamily: adenomatous polyposis coli protein
C:Keywords: cancer; familial adenomatous polyposis; tumor suppressor
F:1-730/Domain: leucine-rich <NTR>
F:7-72/Region: coli #status predicted
F:185-227/Region: coli #status predicted
F:731-2832/Domain: serine-rich <CTD>
F:1131-1156/Region: acidic
F:1558-1577/Region: acidic
F:1866-1893/Region: highly charged
Alignment Scores:
Pred. No.: 39.7 Length: 2843
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 1 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x RBHUAP (1-2843)
QY 959 TTCTCCAGGTGACAGCTTGAGT 936
Db 1271 PheSerArgCysSerSerLeuSer 1278
RESULT 29
149505
adenomatous polyposis coli protein - mouse
N:Alternate names: APC
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
A:Accession: I49505
R:Su, L.
Science 256, 668-670, 1992
A:Title: Multiple intestinal neoplasia caused by a mutation in the murine homolog of th
A:Reference number: I49505; MUID:92263101; PMID:1350108
A:Accession: I49505
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-2845 <RES>
A:Cross-references: GB:M88127; NID:G191991; PIDN:AA59632.1; PID:G191992
C:Superfamily: adenomatous polyposis coli protein
Alignment Scores:
Pred. No.: 39.7 Length: 2845
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 2 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x I49505 (1-2845)
QY 959 TTCTCCAGGTGACAGCTTGAGT 936
Db 1269 PheSerArgCysSerSerLeuSer 1276
RESULT 30
B97073
beta-glucosidase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
A:Accession: B97073
R:Nozling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C.
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B97073
A:Status: preliminary

A/Molecule type: DNA
A/Residues: 1-473 <KUR>
A/Cross-references: GB:AB001437; PIDN:AAK79373.1; PID:G15024344; GSPDB:GN00168
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetics:
A/Gene: CAC1405
C/Superfamily: Agrobacterium beta-glucosidase

Alignment Scores:
Pred. No.: 554 Length: 473
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x B97073 (1-473)

QY 1791 ATGACTTGCTCCGCAAGT 1771
DB 422 lIeAspleuValSerAlaSer 428

RESULT 31
D91075
6-phospho-beta-glucosidase [imported] - Escherichia coli (strain O157:H7, substrain R1MD
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C/Accession: D91075
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kunihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: D91075
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-474 <HAY>
A/Cross-references: GB:BA000007; PIDN:BA836995.1; PID:G13363043; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain R1MD 0509952
C/Genetics:
A/Gene: EC63572
C/Superfamily: Agrobacterium beta-glucosidase

Alignment Scores:
Pred. No.: 554 Length: 474
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x D91075 (1-474)

QY 1791 ATGACTTGCTCCGCAAGT 1771
DB 422 lIeAspleuValSerAlaSer 428

RESULT 32
C85920
6-phospho-beta-glucosidase, cryptic [imported] - Escherichia coli (strain O157:H7, subst
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: C85920
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Iller, L.; Grobleck, E.J.; Davis, N.W.; Lim, A.; Dimmlanta, E.; Potamouets, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: AB5480; MUID:21074935; PMID:11206551
A/Accession: C85920
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-474 <STO>
A/Cross-references: GB:AB005174; NID:g12517164; PIDN:AMG57823.1; GSPDB:GN00145; UWGP:Z40

A/Experimental source: strain O157:H7, substrain EDJ933
C/Genetics:
A/Gene: ascB
C/Superfamily: Agrobacterium beta-glucosidase

Alignment Scores:
Pred. No.: 554 Length: 474
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x H65051 (1-474)

QY 1791 ATGACTTGCTCCGCAAGT 1771
DB 422 lIeAspleuValSerAlaSer 428

RESULT 33
H65051
6-phospho-beta-glucosidase (BC 3.2.1.86) - Escherichia coli (strain K-12)
C/Species: Escherichia coli
C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C/Accession: H65051; C44070; G27553
R/Plattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: H65051
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-474 <BLAT>
A/Cross-references: GB:AE000355; GB:U00096; NID:92367151; PIDN:AAK75758.1; PID:g1789070,
A/Experimental source: strain K-12, substrain MG1655
R/Hall, B.G.; Xu, L.
Mol. Biol. Evol. 9, 688-706, 1992
A/Title: Nucleotide sequence, function, activation, and evolution of the cryptic asc ope
A/Reference number: A44070; MUID:92334140; PMID:1630307
A/Accession: C44070
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-404, 'GT', 407-427, 'C', 429-454, 'HR', 457-474 <HAL>
A/Cross-references: GB:W33326; NID:g145385; PIDN:AAI6430.1; PID:g145388
A/Experimental source: strain LP103
A/Note: sequence extracted from NCBI backbone (NCBI:109109, NCBI:P.109114)
C/Genetics:
A/Gene: ascB
A/Map position: 59 min
C/Superfamily: Agrobacterium beta-glucosidase
C/Keywords: glycosidase; hydrolase

Alignment Scores:
Pred. No.: 554 Length: 474
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x H65051 (1-474)

QY 1791 ATGACTTGCTCCGCAAGT 1771
DB 422 lIeAspleuValSerAlaSer 428

RESULT 34
S49886
probable membrane protein YIL123w - Yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein YI8277.06
C/Species: Saccharomyces cerevisiae
C/Date: 13-Jan-1995 #sequence_revision 20-Feb-1995 #text_change 11-Jan-2000

C/Accession: S49886
R/Hamlyn, N.; Churcher, C.
Submitted to the EMBL Data Library, November 1994
A/Reference number: S49881
A/Accession: S49886
A/Molecule type: DNA
A/Residues: 1-475 <KLE>
A/Cross-references: GB:Z47047; EMBL:Z46833; NID:ig603997; PID:ig763223; MIPS:YLL123W
C/Genetics:
A/Gene: SGD:SIM1
A/Cross-references: SGD:S0001385; MIPS:YLL123W
A/Map position: 9L
C/Superfamily: Saccharomyces NC3 protein
C/Keywords: transmembrane protein
P/82-98/Domain: transmembrane #status predicted <TM>

Alignment Scores:
Pred. No.: 554 Length: 475
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x S49886 (1-475)

QY 1058 CCTGCTGCTCTGAGTAGCA 1038
DB 57 ProAlaAlaSerGluValAla 63

RESULT 35
B70143
fibronectin/fibrinogen-binding protein homolog - Lyme disease spirochete
C/Species: Borrelia burgdorferi (Lyme disease spirochete)
C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C/Accession: B70143
R/Faaser, C.M.; Caslens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugr,
Nature 390, 580-586, 1997
A/Authors: Smith, H.O.; Venter, J.C.
A/Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A/Reference number: A70100; MUID:98065943; PMID:9403685
A/Accession: B70143
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-475 <KLE>
A/Cross-references: GB:AE001141; GB:AE000783; NID:G2688250; PIDN:AA06734.1; PID:G268825
A/Experimental source: strain B31

Alignment Scores:
Pred. No.: 554 Length: 475
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.05% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x B70143 (1-475)

QY 892 AAGAGTTAAGACCGACT 912
DB 469 Lysblyllylsbnglntr 475

RESULT 36
H89773
hypothetical protein SA0127 [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C/Accession: H89773
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mitani, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; MUID:21311952; PMID:11418146
A/Accession: H89773
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-476 <KUR>
A/Cross-references: GB:BA000018; PID:g13700048; PIDN:BA041347.1; GSPDB:GN00149
A/Experimental source: strain N315
C/Genetics:
A/Gene: SA0127

Alignment Scores:
Pred. No.: 554 Length: 476
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.05% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x H89773 (1-476)

QY 452 AGCCTTGCTTACCTGTGCCA 472
DB 211 SerLeuAlaTyrLeuLeuPro 217

RESULT 37
E97073
phospho-beta-glucosidase [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C/Accession: E97073
R/Molling, U.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium CJ
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: E97073
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-477 <KUR>
A/Cross-references: GB:AE001437; PIDN:AAK79376.1; PID:G15024347; GSPDB:GN00168
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetics:
A/Gene: CA01408
C/Superfamily: Agrobacterium beta-glucosidase

Alignment Scores:
Pred. No.: 553 Length: 477
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x E97073 (1-477)

QY 1791 ATGACTGTGCTCCGCAAGT 1771
DB 425 IleAspLeuValSerAlaSer 431

RESULT 38
G83048
probable two-component sensor PA4777 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C/Accession: G83048
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kass, A.; Lapidig, K.; Lm
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A/Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: G83048
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-477 <STO>
A:Cross-references: GB:AE004891; GB:AE004091; NID:g9951037; PIDN:AA608163.1; GSPDB:GN001
A:Experimental source: strain PMOI
C:Genetics:
A:Gene: PA4777

Alignment Scores:

Pred. No.:	553	Length:	477
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x G83048 (1-477)

QY 1841 CTCGCGCGAGTCAGCTCTG 1821

Db 90 LeuArgArgGlnSerAlaLeu 96

RESULT 39

beta-glucosidase (EC 3.2.1.21) [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1)
C:Species: *Lactococcus lactis* subsp. *lactis*
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: C86807
R:Boletín, A.; Winkler, P.; Manger, S.; Jallón, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s8
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: C86807
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-478 <STO>
A:Cross-references: GB:AE005176; PID:g12724451; PIDN:AAK05557.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: bgjH
C:Superfamily: Agrobacterium beta-glucosidase
C:Keywords: glycosidase; hydrolase

Alignment Scores:

Pred. No.:	553	Length:	478
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x C86807 (1-478)

QY 1791 ATAGACTGGTCTCCGCAAGT 1771

Db 427 IleAspLeuValSerAlaSer 433

RESULT 40

S17836
fatty-acyl-CoA reductase (EC 1.2.1.-) luxC - *Photobacterium leiognathi*
C:Species: *Photobacterium leiognathi*
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 26-May-2000
C:Accession: S17836
R:Lee, C.Y.; Sziltemer, R.B.; Weighen, E.A.
Eur. J. Biochem. 201, 161-167, 1991
A>Title: The lux genes of the luminous bacterial symbiont, *Photobacterium leiognathi*, of
coll.
A:Reference number: S17836; MUID:92007870; PMID:1915359
A:Accession: S17836
A:Molecule type: DNA
A:Residues: 1-478 <LEB>
A:Cross-references: EMBL:M63594; NID:g150687; PIDN:AAA25616.1; PID:g150688

C:Genetics:
A:Gene: luxC
C:Superfamily: Photobacterium fatty-acyl-CoA reductase
C:Keywords: oxidoreductase

Alignment Scores:

Pred. No.:	553	Length:	478
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.05%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x S17836 (1-478)

QY 355 TCTGTTACCTGCTTGCAAGT 375

Db 145 SerValHisLeuValAlaGly 151

RESULT 41

F70573
hypothetical protein RV0341 - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: F70573
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
Rajandream, M.A.; Rogers, R.; Devlin, K.; Felkwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.
Nature 393, 537-544, 1998
A:Author: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70573
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-479 <COL>
A:Cross-references: GB:Z95324; GB:AL123456; NID:g3261760; PIDN:CAB08573.1; PID:g27094820
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0341
C:Superfamily: Phaeolus glycine-rich cell wall protein 1.8

Alignment Scores:

Pred. No.:	553	Length:	479
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x F70573 (1-479)

QY 1491 GCGGTGCCACTGCTGCTGGT 1471

Db 342 GlycylAlaThrAlaIaIaGly 348

RESULT 42

T47415
transporter-like protein - *Arabidopsis thaliana*
N:Alternate names: protein T28A8.80
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47415
R:Purnelle, B.; Boutry, M.; Goffeau, A.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24466
A:Accession: T47415
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <PUR>
A:Cross-references: EMBL:AL162691
A:Experimental source: cultivar Columbia; BAC clone T28A8
C:Genetics:

A:Map position: 3
A:introns: 48/1: 77/1; 111/2; 145/3; 163/3; 184/3; 201/2; 226/3; 291/3; 321/1; 349/3; 38
A:Note: T28A8.80

Alignment Scores:
Pred. No.: 553 Length: 479
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.05% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x T47415 (1-479)

QY 1558 CTTTGGCCTTTCACACAC 1578
|||||
Db 117 Leupheg1yenseTher 123

RESULT 43
H86411
Protein PK23.12 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H86411
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
amem, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
Ker, M.; Wu, D.; Yu, G.; Fraser, C.W.; Venter, J.C.; Davis, R.W.
A:title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H86411
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-480 <STO>
A:Cross-references: GB:AE005172; NID:G10764863; PIDN:AAF24555.2; GSPDB:GN00141
C:Genetics:
A:Gene: PK23.12
A:Map position: 1

Alignment Scores:
Pred. No.: 553 Length: 480
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x H86411 (1-480)

QY 57 AGCACCACTCTCAGTAC 37
|||||
Db 15 SerHidg1yenseTher 21

RESULT 44
H69593
6-phospho-beta-glucosidase (EC 3.2.1.86) bg1H [similarity] - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C:Accession: H69593; I40407; T47098; S47175
R:Kunert, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berti,
C.; Bron, S.; Broutillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holappel, S.; Hosono, S.; Hullo, M.F.;
Koester, P.; Koningsreith, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
A:Authors: Schlach, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowski, A.; Sero
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Tappera, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: H69593
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-481 <KUN>
A:Cross-references: GB:Z99124; GB:AL009126; NID:G2636442; PIDN:CAB15962.1; PID:G2636472
A:Experimental source: strain 168
R:le Cog, D.; Lindner, C.; Kruger, S.; Steinmetz, M.; Stulke, J.
U. Bacteriol. 177, 1527-1535, 1995
A:title: New beta-glucosidase (bg1) genes in Bacillus subtilis: the bg1 gene product has
A:reference number: I40404; MUID:95189730; PMID:7883710
A:Accession: I40407
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 13-413, R', 415-481 <LEC>
A:Cross-references: EMBL:Z34526; NID:G505573; PIDN:CAA84287.1; PID:G505577
R:Yoshida, K.; Sano, H.; Seki, S.; Oda, M.; Fujimura, M.; Fujita, Y.
Microbiology 141, 337-343, 1995
A:title: Cloning and sequencing of a 29 kb region of the Bacillus subtilis genome conta
A:reference number: Z24350; MUID:95219088; PMID:7704263
A:Accession: I47098
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-481 <ROS>
A:Cross-references: EMBL:D31856; NID:G603765; PIDN:BAA06653.1; PID:G603779
A:Experimental source: strain BGSC1A1
C:Genetics:
A:Gene: bg1H; M7D
C:Superfamily: Agrobacterium beta-glucosidase
C:Keywords: glucosidase; hydrolase

Alignment Scores:
Pred. No.: 553 Length: 481
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x H69593 (1-481)

QY 1791 ATGACTGTCTCGCAGT 1771
|||||
Db 430 Ileapleuvalseralaser 436

RESULT 45
AB1421
beta-glucosidase homolog lmo2771 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AB1421
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Anand, A.; Baquero, F.; Berche, P.; Blocke,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Feihl, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A:Authors: Krefel, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; M.
Ok, C.; Schluter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,
A:title: Comparative genomics of Listeria species.
A:reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1421
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-482 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00984.1; PID:G16412271; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2771

C:/Superfamily: Agrobacterium beta-glucosidase

Alignment Scores:

Pred. No.:	552	Length:	482
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AB1421 (1-482)

QY 1791 ATGAGCTGTCTCCGCAAGT 1771

Db 428 IleaSpneuValSerAlaSer 434

RESULT 46

A31521

Orphan receptor TR2, splice form TR2-5 - human

C/Species: Homo sapiens (man)

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000

C/Accession: A31521

R/Chang, C.; Kokontis, J.

Biochem. Biophys. Res. Commun. 155, 971-977, 1988

A/Title: Identification of a new member of the steroid receptor super-family by cloning

A/Reference number: A31521; MUID:88339993; PMID:3421977

A/Accession: A31521

A/Molecule type: mRNA

A/Residues: 1-483 <CH>

A/Cross-references: GB:M21985; NID:g338485; PIDN:AAA3650.1; PID:g338486

C/Genetics:

A/Gene: GDB:TR2-11

A/Cross-references: GDB:5875387

C/Superfamily: orphan receptor TR2; erba transforming protein homology

C/Keywords: alternative splicing; DNA binding; nucleus; receptor; transcription regulati

F/111-465/Domain: erba transforming protein homology #status atypical <ERBA>

F/113-133/Region: zinc finger CCCC motif

F/149-168/Region: zinc finger CCCC motif

Alignment Scores:

Pred. No.:	552	Length:	483
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	1	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x A31521 (1-483)

QY 487 GGGTGGCCCAATGTTGGCA 467

Db 437 gLyLeuAlaGlnCysTrpGln 443

RESULT 47

F87700

rfaE protein [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C/Accession: F87700

R/Nierman, W.C.; Felldjlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolenko, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, D.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of Caulobacter crescentus.

A/Reference number: AB7249; MUID:21173698; PMID:11259647

A/Accession: F87700

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-483 <STO>

A/Cross-references: GB:AE005673; NID:g13425394; PIDN:AAK25602.1; GSPDB:GN00148

C/Genetics:

A/Gene: CC3640

C/Superfamily: hypothetical protein b3052

Alignment Scores:

Pred. No.:	552	Length:	483
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.05%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x F87700 (1-483)

QY 1829 ACTGGCGGCGAGCTGGCATG 1849

Db 239 ThrArgGlyGlyAlaGlyMet 245

RESULT 48

B33501

myosin heavy chain 2, smooth muscle - rabbit (fragment)

C/Species: Oryctolagus cuniculus (domestic rabbit)

C/Date: 08-Dec-1989 #sequence_revision 08-Dec-1989 #text_change 13-Aug-1999

C/Accession: B33501; A31368

R/Nagel, R.; Kuro-o, M.; Babij, P.; Periasamy, M.

J. Biol. Chem. 264, 9734-9737, 1989

A/Title: Identification of two types of smooth muscle myosin heavy chain isoforms by cDN

A/Reference number: A33501; MUID:89255535; PMID:2722872

A/Accession: B33501

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-484 <NAG>

A/Cross-references: GB:J04833

R/Nagel, R.; Larson, D.M.; Periasamy, M.

Proc. Natl. Acad. Sci. U.S.A. 85, 1047-1051, 1988

A/Title: Characterization of a mammalian smooth muscle myosin heavy chain cDNA clone an

A/Reference number: A31368; MUID:88124972; PMID:3422477

A/Accession: A31368

A/Molecule type: mRNA

A/Residues: 1-484 <NAG>

A/Cross-references: GB:J03614; NID:g165517; PIDN:AAA31406.1; PID:g165518

C/Superfamily: myosin heavy chain; myosin motor domain homology

C/Keywords: muscle; smooth muscle

Pred. No.:	552	Length:	484
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x B33501 (1-484)

QY 1403 CTGGCTCCCTTCATCAGCA 1383

Db 263 LeuAlaSerSerLeuSerGly 269

RESULT 49

AF1107

phospho-beta-glucosidase homolog lmo0261 [imported] - Listeria monocytogenes (strain BGD

C/Species: Listeria monocytogenes

C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C/Accession: AF1107

R/Glaser, P.; Farnagel, L.; Buchrieser, C.; Amend, A.; Baguerio, F.; Berche, P.; Bloeker

D.; Dominguez-Jernai, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Feint, H.

Science 294, 849-852, 2001

A/Authors: Krefl, J.; Kunz, M.; Kunz, F.; Kurapat, G.; Macleno, E.; Maltournam, A.; Ma

ok, C.; Schluter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,

A/Title: Comparative genomics of Listeria species.

A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AF1107

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-486 <GUA>

A:Cross-references: GB:NC_003210; PIDN:CAD00788.1; PID:gl6409626; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo0261
 C:Superfamily: Agrobacterium beta-glucosidase

Alignment Scores:
 Pred. No.: 552 Length: 486
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AP1107 (1-486)

QY 1791 ATGACTTGCTCGCGAAGT 1771

DB 434 lIeApIeuValSerAlaser 440

RESULT 50

AT1468
 phospho-beta-glucosidase homolog lin0288 [imported] - Listeria innocua (strain Clp11262
 C:Species: Listeria innocua
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #ext_change 14-Dec-2001
 C:Accession: AT1468
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, U.; Kuhn, M.; Kunz, F.; Kurapkai, G.; Madueno, E.; Maltournan, A.; Ma
 ok, C.; Schlueter, T.; Simoes, N.; Tlerez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlard,
 A.:Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; PMID:11537279; PMID:11679669
 A:Accession: AT1468
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-486 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC95521.1; PID:gl6412717; GSPDB:GN00178
 A:Experimental source: strain Clp11262
 C:Genetics:
 A:Gene: lin0288
 C:Superfamily: Agrobacterium beta-glucosidase

Alignment Scores:
 Pred. No.: 552 Length: 486
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AT1468 (1-486)

QY 1791 ATGACTTGCTCGCGAAGT 1771

DB 434 lIeApIeuValSerAlaser 440

Search completed: June 21, 2004, 13:15:16
 Job time : 103.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:46:21 ; Search time 22 Seconds
(without alignments)
9472.035 Million cell updates/sec

Title: US-10-029-345A-108_COPY_532_2532
Perfect score: 667
Sequence: 1 gttcgtcatgcccacatgagat.....tggaaatcatgaggtctcc 2001

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Word size: 1

Total number of hits satisfying chosen parameters: 66812

Minimum DB seq length: 473
Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10029345/runat_21062004_124615_5680/app_query.fasta_1.2183
-DB=SwissProt 42 -QFMT=fasta -SUFFIX=oligo.rsp -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=oligo -TRANS=human0.cdi
-LIST=75 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL
-OUTFMT=proto -NORM=exc -HEAPSIZE=500 -MINLEN=473 -MAXLEN=200000000
-USER=US10029345_@CGN_1_1_22 @runat_21062004_124615_5680 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEXT TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCAPOP=6
-FCAPEXT=7 -YCAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB	ID
1	472	70.8	665 1 DUSG_HUMAN
2	20	3.0	625 1 DUS8_HUMAN
3	20	3.0	663 1 DUS8_MOUSE
4	9	1.3	619 1 VHP1_CABEL
5	9	1.4	737 1 ACET_RABIT
6	9	1.4	890 1 MUTS_RICCN
7	8	1.2	480 1 HBM3_EUGAR
8	8	1.2	509 1 AGP4_MOUSE
9	8	1.2	526 1 VPS_BTVO
10	8	1.2	555 1 VGLF_PIIHC
11	8	1.2	605 1 GOX_ASPNG
12	8	1.2	640 1 UH35_HCVNA
13	8	1.2	677 1 HUTH_CABEL
14	8	1.2	835 1 VIRL_AGRU
15	8	1.2	1716 1 RPAL_RAT
16	8	1.2	1717 1 RPAL_HUMAN
17	8	1.2	1717 1 RPAL_MOUSE
18	8	1.2	2512 1 POLN_SINDV

DESCRIPTION

Q9BY84 homo sapien
Q13202 homo sapien
O09112 mus musculi
Q10038 caenorhabdi
P22968 oryctolagus
O92119 rickettsia
P13446 euglena gra
Q9WV16 mus musculi
P07389 blueongue
P12605 aspergillus
P13006 human parai
P16766 human cytom
Q20502 caenorhabdi
P07167 agrobacteri
O54889 rattus norv
O95602 homo sapien
O35134 mus musculi
P03317 sindbis vir

ALIGNMENTS

RESULT 1
DUSG_HUMAN
ID DUSG_HUMAN STANDARD; PRT; 665 AA.
AC Q9BY84; Q9C0G3;
DT 28-FEB-2003 (Rel. 41, last sequence update)
DT 28-FEB-2003 (Rel. 41, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE Dual specificity protein phosphatase 16 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 7) (MAP kinase
DE phosphatase 7) (MKP-7).
GN DUSP16 OR MKP7 OR KIAA1700.
OS Homo sapiens (human).

19 8 1.2 2514 1 POLN_SINDV
20 8 1.2 2842 1 APC_RAT
21 8 1.2 2842 1 APC_HUMAN
22 8 1.2 2845 1 APC_MOUSE
23 8 1.1 473 1 ABGA_CLOLO
24 7 1.0 473 1 ATPB_PRES
25 7 1.1 474 1 ASCB_ECOLI
26 7 1.0 475 1 PERI_MOUSE
27 7 1.1 475 1 SIM1_YEAST
28 7 1.0 478 1 LXC2_PHOIE
29 7 1.1 483 1 TR2_HUMAN
30 7 1.0 487 1 YMI8_STRAW
31 7 1.1 489 1 VGLY_MOPEI
32 7 1.1 491 1 FLIC_YEAST
33 7 1.1 494 1 FLIC_SALTY
34 7 1.1 494 1 RNPL_HUMAN
35 7 1.1 496 1 HUTH_TREAC
36 7 1.0 503 1 MATK_LIPO
37 7 1.0 503 1 MATK_LIGOR
38 7 1.0 503 1 MATK_LIGOR
39 7 1.0 507 1 MS14_ARATH
40 7 1.0 510 1 PUR1_CHICK
41 7 1.0 511 1 HUTH_VIBCH
42 7 1.1 517 1 COX1_MYGGL
43 7 1.0 518 1 YH99_YEAST
44 7 1.1 521 1 YAVB_SCHPO
45 7 1.0 524 1 HEMK_RICCN
46 7 1.1 526 1 CAP_YEAST
47 7 1.0 532 1 PPB_SCHPO
48 7 1.0 535 1 TCBE1_AVEA
49 7 1.0 535 1 TCBE2_AVEA
50 7 1.0 535 1 TCPE2_ARATH
51 7 1.1 540 1 CH60_THERB
52 7 1.1 545 1 PYRG_VIBVU
53 7 1.1 548 1 CH60_OLEAN
54 7 1.1 548 1 TCPEQ_HUMAN
55 7 1.1 548 1 TCPEQ_MOUSE
56 7 1.1 550 1 WR27_CABEL
57 7 1.1 551 1 CEA3_ECOLI
58 7 1.0 559 1 GL65_CAPII
59 7 1.0 564 1 NOX1_HUMAN
60 7 1.0 566 1 SCCL_YEAST
61 7 1.1 567 1 FIG1_HUMAN
62 7 1.0 567 1 IF37_SCHPO
63 7 1.0 574 1 CIRS_HUMAN
64 7 1.1 577 1 YDEM_RHIME
65 7 1.1 578 1 MALR_AEDAE
66 7 1.0 579 1 SCB3_SCHPO
67 7 1.0 583 1 ACN3_CABEL
68 7 1.1 585 1 ILV3_YEAST
69 7 1.0 585 1 ILV3_YEAST
70 7 1.1 588 1 ATY1_MOUSE
71 7 1.0 591 1 UL49_EBV
72 7 1.0 595 1 SIL1_HUMAN
73 7 1.1 596 1 CHVG_AGRIS
74 7 1.0 597 1 SIL1_PANTR
75 7 1.0 597 1 Y745_ARATH

P27283 sindbis vir
P70478 rattus norv
P25054 homo sapien
O61315 mus musculi
Q46130 clostridium
O03080 pteridium e
P24240 escherichia
P15331 mus musculi
P40499 saccharomyc
P29236 photoactet
P13056 homo sapien
Q82K24 streptomyce
P19240 mopeia viru
P40499 saccharomyc
P06179 salmonella
Q9HAU8 homo sapien
Q9H116 thermoplasma
Q98370 liqidambar
O8W4M5 liqidambar
Q99073 liqidambar
Q22607 arabidopsis
P28173 gallus gall
Q9KEG4 vibrio chol
O21079 myxine glut
P38758 saccharomyc
Q10177 schizosacch
Q92913 rickettsia
P17555 saccharomyc
O60109 schizosacch
P40412 avena sativ
P54411 avena sativ
O04450 arabidopsis
O60024 thermoaer
Q8DE33 vibrio vuln
Q8KM30 oleispirita a
P50390 homo sapien
P42932 mus musculi
Q94130 caenorhabdi
P00646 escherichia
P50426 capra hircu
Q9Y588 homo sapien
Q12158 saccharomyc
Q96929 homo sapien
O94236 schizosacch
O43734 homo sapien
P72292 thizobium m
Q19124 caenorhabdi
P13080 aedes aegypt
P1000 schizosacch
Q9U785 caenorhabdi
P39522 saccharomyc
Q9CE66 mus musculi
P14347 epstein-bar
Q96921 homo sapien
O07737 agrobacteri
Q95180 pan troglod
P59278 arabidopsis

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21486429; PubMed=11489891;
 RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
 RT "MKP-7, a novel mitogen-activated protein kinase phosphatase,
 RT functions as a shuttle protein";
 RL J. Biol. Chem. 276:39002-39011(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=21082932; PubMed=11214970;
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIX.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:347-355(2000).
 CC -1 FUNCTION: Involved in the inactivation of MAP kinases.
 CC -1 CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate
 CC -1 CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
 CC -1 SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1 SIMILARITY: Contains 1 rhodanese domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB052156; BAB40814.1; -
 DR EMBL; AB051487; BAB21791.1; ALT_INIT.
 DR HSSP; Q16828; IMKP.
 DR Genew; HGNC:17909; DUSP16.
 DR MIM; 607175; -
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0004721; F:protein phosphatase activity; TAS.
 DR GO; GO:0016311; F:dephosphorylation; TAS.
 DR GO; GO:0000188; P:inactivation of MAPK; TAS.
 DR GO; GO:0045209; P:leptomycin B-sensitive MAPK phosphatase nuc. .; TAS.
 DR GO; GO:0045204; P:MAPK nucleus export; TAS.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPc; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PRO1764; MAPKPHPTASE.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS02026; RHODANES_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
 KM Hydrolyase; Nucleic acid protein.
 FT DOMAIN 22 137 RHODANES.
 FT DOMAIN 227 289 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 244 244 PHOSPHOCYSTEINE INTERMEDIATE (BY
 FT SIMILARITY)
 SQ SEQUENCE 665 AA; 73101 MW; 1B0853FF08460DFP CRC64;

Alignment Scores:
 Pred. No.: 0
 Score: 472.00
 Percent Similarity: 100.00%

Length: 665
 Matches: 472
 Conservative: 0

Best Local Similarity: 100.00%
 Query Match: 70.76%
 DB: 1
 Gaps: 0
 US-10-029-345a-108_COPY_532_2532 (1-2001) x DUSG_HUMAN (1-665)
 QY 586 ACCTGTCGAAGAGCTGACTTATATCCCGAGTCTCATTTCTCGTGGTGGTGAATGAC 645
 |||||
 DB 194 TtrCysProlyserProspheilleleprogiuserhisleuvalprovalasasp 213
 QY 646 AGCTTTGAGAGAAATTTGCGGTGGTGGCAATGATGATTGATGAGAAACA 705
 |||||
 DB 214 SerpneCysgluylleuProtrpueasplyservalaspheileglulysala 233
 QY 706 AAAGCTCCAAATGATGATGTTTCTAGTGCATGTTAGCTGGAGATCCCGCTCCGCAAC 765
 |||||
 DB 234 LysAlaserAanglYcysvalleuvalHisCysleuAlaaglylleSerArgserlaThr 253
 QY 766 ATCCGCTATGCTACATCATGAGAGATGACATGCTTTAGATGAAGCTTACGATTT 825
 |||||
 DB 254 llealallealatyrllewetlysargmetserleuaspgluallatyrrpHe 273
 QY 826 GTGAAGAAAAAGACCTACTATATCTCAACTCATTTTGGGGCAACTCTGGAC 885
 |||||
 DB 274 VallysgluysargprothrilleSerProasnpheasplleuGlylnleuasp 293
 QY 886 TATGAAGAAAGATTAGAACCAAGCTGAGCATGAGGCCCAAGAGCAACTGACGTG 945
 |||||
 DB 294 TyrGluylsylelleYasnglnthrGlylaserGlyprolysersyleuylsleu 313
 QY 946 CTGCACTGGAGAGGCAAAAGAACTGTCCTGCTGTCTAGAGGTGGACAGAAAGC 1005
 |||||
 DB 314 LeuHlsleuGluysProasngluProvalProalvalaserGluGlyGlnlyser 333
 QY 1006 GAGAGCCCTCAGTCCACCTGTGCGCACTCTGCTCACTCAGAGGAGAGCAAGAG 1065
 |||||
 DB 334 GlutHrProleuserProProCysAlaaspserrlaThrserGluAlaaglylnHrg 353
 QY 1066 CCCGTGCATCCCGCCGAGGTGCCAGGCGTCCAGCGTGCAGCGCTGTTAGAGAC 1125
 |||||
 DB 354 ProvalHsProalaserValProserValProserValGlnProserleuGluasp 373
 QY 1126 AGCCGCTGTACAGGCGCTCAGTGGGCTGCACTGTCGGCAGACAGCTGGAGAGAC 1185
 |||||
 DB 374 SerProleuvalGlnAlaaleuSerGlyleuHlsleuSerAlaspaGluaspSer 393
 QY 1186 AATAAGCTCAAGCGTTCCTCTCTCTGATATCAATCAGTTTCATATTCAGCAGCATG 1245
 |||||
 DB 394 AsnlyseuylsargserPheSerleuaspilleyservalserTyrserAlaserMet 413
 QY 1246 GCACATCTTACATGAGCTTCTCTCATCGAAGATGCTTTGGAATTCACAACTTCC 1305
 |||||
 DB 414 AlaAlaserleuHlsGlyPheSerSerSerGluaspAlaaleuGluTyrTylsProser 433
 QY 1306 ACTACTCTGATGGAGCAACAAGCTATGCACTGTCAGTTCCTCCCTGTCAGCAATATCGAG 1365
 |||||
 DB 434 ThrThrleuaspGlylnHlsleuylsleuGlylnPheSerProvalGlnGluSerGln 453
 QY 1366 CAGACTCCGAAACAGCTCTGTATTAAGAGAAAGCAGCATCTCCCAAGAGTGCAGACC 1425
 |||||
 DB 454 GlnThrProgluThrserProasplyGluGluAlaserilleProlysyleuGlnThr 473
 QY 1426 GCGAGCGCTTCAGACGCCAGCAGCAAGCATTCGATTCGGTCAAGACAGCAGCATGGC 1485
 |||||
 DB 474 AlaArgProserAspserGlnserlYsargleuHlservalArgThrserSerSerGly 493
 QY 1486 ACCGCCAGAGGTCCCTTTATCTCCAGTCATCAAGTGGAGCGGTGGAGCAATTAC 1545
 |||||
 DB 494 ThrAlaGlnArgserleuSerProleuHlsargserGlyserValGluaspaentYr 513
 QY 1546 CACACCAAGCTCTTTTGGCTTTGCACAGCAGCAGCAGCAGCAGTGTGCTGAC 1605
 |||||
 DB 514 HisThrserPheleuPheGlyleuSerThrSerGlnGlnHlsleuThrlysserAlaGly 533


```

QY 1606 CTGGGCTTAAGGGCTGGGCACTGGATATCTGGCCCCCAGACCTTACCCCTTCCG 1665
DB 534 LeuGlyLeuLeuSerIYrPhaSerAapIleLeuAlaProIlnThrSerThrProSerLeu 553
QY 1666 ACCAGCAGCTGTATTTTGGCCAGAGTCTCTCACTTCTACTCTGCTCAGCCATCTAC 1725
DB 554 ThSerSerTrpYrPhaThrAlaThrGlnSerSerHisPheYrSerIlaSerAlaIleYr 573
QY 1726 GAGGAGCTGCCGATTACTCTGCTTACAGCTGACGAGCCAGCTGCCACTTGGGAGACCAA 1785
DB 574 GlyGlySerAlaSerIYrSerAlaYrSerCysSerGlnLeuProThrCysGlyAapGln 593
QY 1786 GTCTATTTCTGGGCGGCGGCGGAGAGCCAGAGCTGACCTCGGCGGCGGAGCTCG 1845
DB 594 ValIYrSerValAlaYrAlaGlnYrGlnYrProSerAapYrAlaApsSerAlaYrSerTrp 613
QY 1846 CATGAGAGAGAGCCCTTTGAAAAGCAGTTTAAACGAGAGAGCTGCGCAATGGAATTGG 1905
DB 614 HisGlnGlnSerProPheGlnYrGlnPheYrAlaYrGlnSerCysGlnMetGlnPheGly 633
QY 1906 GAGAGCATCATGTCTCAAGAAACAGGTCACGGGAAAGCTGGGAAAGTGGGCACTGCT 1965
DB 634 GluSerIleMetSerGlnAenAArgSerArgGlnGlnLeuGlyValGlySerGlnSer 653
QY 1966 AGCTTTTGGGCGAGCATGGAATCATGAGCTCTCC 2001
DB 654 SerPheSerGlySerMetGlnIleIleGlnValSer 665

RESULT 2
DUS8_HUMAN STANDARD; PRT; 625 AA.
AC Q13202;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Dual specificity protein phosphatase hvh-5).
GN DUSP8 OR VHS.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Fetal brain;
RX MEDLINE=60009533; PubMed=7561881;
RA Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.;
RT "hvh-5: a protein tyrosine phosphatase abundant in brain that
RT inactivates mitogen-activated protein kinase.";
RL J. Neurochem. 65:1823-1833(1995).
CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
CC phosphatase as well as with serine/threonine-protein phosphatase (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- TISSUE SPECIFICITY: Abundant in brain, heart and skeletal muscle.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC -1- Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U27193; AAA83151.1; -
CC DR HSSP: Q16826; 1MKP.

```

```

DR Genew; HGNC:3074; DUSP8.
DR MIM: 602038; -.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; TAS.
DR GO; GO:0004188; F:protein activation of MAPK; TAS.
DR GO; GO:0004470; P:protein amino acid dephosphorylation; TAS.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00450; RHOD; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE_3; 1.
DR PROSITE; PS50383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR KMW; KMW; Nuclear protein.
FT DOMAIN 23 138 RHODANSE.
FT DOMAIN 162 430 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 310 550 PRO-RICH.
FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
FT FT SIMILARITY).
SQ SEQUENCE 625 AA; 65840 MW; DCBA14487219666 CRC64;

Alignment Scores:
Pred. No.: 1,13e-11 Length: 625
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x DUS8_HUMAN (1-625)
QY 730 GTGGACATGTTTGGCTGGATCTCCGCTCGGACCATGCTATCGCTACATGATGAG 789
DB 244 ValHisCysLeuAlaIYrIleSerArgSerAlaThrIleAlaIleAlaYrIleMetLys 263

RESULT 3
DUS8_MOUSE STANDARD; PRT; 663 AA.
AC O09112;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Neuronal tyrosine threonine phosphatase 1).
GN DUSP8 OR NTTP1.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=96311565; PubMed=8733137;
RX Theodosiou A.M., Rodrigues N.R., Nesbitt M.A., Ambrose H.J.,
RX Paterson H., McEellan-Arnold E., Boyd Y., Leverhwa M.A., Owen N.,
RX Blake D.J., Ashworth A., Davies K.E.;
RT "A member of the MAP kinase phosphatase gene family in mouse
RT containing a complex trinucleotide repeat in the coding region.";
RL Hum. Mol. Genet. 5:675-684(1996).
CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
CC phosphatase as well as with serine/threonine-protein phosphatase (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- TISSUE SPECIFICITY: Expressed predominantly in brain and lung.

```

CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X95518; CA64772.1; -
CC HSSP; Q16828; IMKP.
CC MGD; MGI:10626; Dusp8.
CC InterPro; IPR000340; DS_phosphatase.
CC InterPro; IPR008343; MAPK_phosph.
CC InterPro; IPR001763; Rhodanese-like.
CC InterPro; IPR000387; TYR_phosphatase.
CC Pfam; PF00782; DSPC; 1.
CC PRINTS; PR01764; MAPKPHPTASE.
CC SMART; SM00195; DSPC; 1.
CC SMART; SM00450; RHOD; 1.
CC PROSITE; PS50206; RHODANES_3; 1.
CC PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
CC PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
CC PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
CC K0 Hydrolyase; Nuclear protein.
CC FT DOMAIN 23 138 RHODANES.
FT DOMAIN 162 432 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 452 459 POLY-ARG.
FT DOMAIN 555 558 POLY-SER.
FT DOMAIN 559 576 POLY-SER.
FT DOMAIN 577 600 POLY-GLY.
FT DOMAIN 311 552 PRO-RICH.
FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
SQ SEQUENCE 663 AA; 68847 MW; 416f429a12c1fa7c CRC64;
SIMILARITY).
Alignment Scores:
Pred. No.: 1,12e-11 Length: 663
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 1 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x DUS8_MOUSE (1-663)
QY 730 GTGCACTGTTTACCTGGATCTCCGCTCCGACCATGCTATGCGCTACATCATGAAG 789
DB 244 VALHICysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 263
RESULT 4
VHP1_CAEEL
ID VHP1_CAEEL STANDARD; PRT; 619 AA.
AC 010038;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein-tyrosine phosphatase vhp-1 (EC 3.1.3.48).
GN VHP-1 OR F08B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Felodoridae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Chisoe S.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.

CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U23178; AAC46719.1; -
CC PIR; T15969; T15969.
CC HSSP; Q16828; IMKP.
CC WormPep; F08B1.1; CE01899.
CC InterPro; IPR000340; DS_phosphatase.
CC InterPro; IPR000387; TYR_phosphatase.
CC Pfam; PF00782; DSPC; 1.
CC SMART; SM00195; DSPC; 1.
CC PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
CC PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
CC PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
CC K0 Hydrolyase.
CC FT DOMAIN 200 269 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 92 95 POLY-SER.
FT DOMAIN 351 354 POLY-SER.
FT DOMAIN 465 472 POLY-SER.
FT DOMAIN 483 488 POLY-SER.
FT ACT_SITE 224 224 PHOSPHOCYSTEINE INTERMEDIATE (BY
SQ SEQUENCE 619 AA; 66354 MW; 369E326F615D0529 CRC64;
SIMILARITY).
Alignment Scores:
Pred. No.: 2.61 Length: 619
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 1 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x VHP1_CAEEL (1-619)
QY 733 CACTGTTTACCTGGATCTCCGCTCCGCTCC 759
DB 223 HICysLeuAlaGlyIleSerArgSer 231
RESULT 5
ACET_RABIT
ID ACET_RABIT STANDARD; PRT; 737 AA.
AC P22968;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Angiotensin-converting enzyme, testis-specific isoform precursor
DE (EC 3.4.15.1) (ACE-T) (Dipeptidyl carboxypeptidase I) (kininase II).
GN ACE OR DCP1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Testis;
RX MEDLINE=89380303; PubMed=2550457;
RA Kumar R.S., Kueari J., Roy S.N., Soffer R.L., Sen G.C.;
RT "Structure of testicular angiotensin-converting enzyme. A segmental
RT mosaic isozyme.";
RL J. Biol. Chem. 264:16754-16758 (1989).
RN [2]
RP SEQUENCE OF 1-117 FROM N.A.
RC MEDLINE=91139683; PubMed=1847388;
RA Kumar R.S., Thekkumkara T.J., Sen G.C.;
RT "The mRNAs encoding the two angiotensin-converting isozymes are

transcribed from the same gene by a tissue-specific choice of
RT alternative transcription initiation sites.";
RL J. Biol. Chem. 266:3854-3862(1991).
CC -1- FUNCTION: Converts angiotensin I to angiotensin II by release of
CC the terminal His-Leu, this results in an increase of the
CC vasoconstrictor activity of angiotensin.
CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide-|-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
CC Asp nor Glu. Converts angiotensin I to angiotensin II.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Testis-specific;
CC IsoId=P22968-1; Sequence=Displayed;
CC Name=Somatic;
CC IsoId=P12822-1; Sequence=External;
CC -1- TISSUE SPECIFICITY: Spermatoocytes, adult testis.
CC -1- INDUCTION: Expression is thought to be subject to hormonal
CC regulation by androgens.
CC -1- SIMILARITY: Belongs to peptidase family M2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05041; AAA31153.1; -;
DR EMBL; M58580; AAA31152.1; -;
DR PIR; A34402; A34402.
DR MEROPS; M02.004; -;
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR001548; Peptidase_M2.
DR Pfam; PF01401; Peptidase_M2; 1.
DR PRINTS; PR00791; PEPTIDASEA.
DR ProDom; PD004184; Peptidase_M2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Glycoprotein;
KW Transmembrane; Testis; Signal; Alternative splicing.
FT CHAIN 1 32
FT SIGNAL 1 32
FT 33 737
FT 33 737
FT DOMAIN 33 690
FT TRANSMEM 691 707
FT DOMAIN 708 737
FT METAL 419 419
FT ACT SITE 420 420
FT METAL 423 423
FT METAL 447 447
FT DISULFID 188 194
FT DISULFID 388 406
FT DISULFID 574 586
FT CARBOHYD 108 108
FT CARBOHYD 126 126
FT CARBOHYD 145 145
FT CARBOHYD 373 373
FT CARBOHYD 622 622
SQ SEQUENCE 737 AA; 83923 MW; FC43CCTG655C3DCA CRC64;

Alignment Scores:
Pred. No.: 2.56 Length: 737
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: 1 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x ACET_RABIT (1-737)
QY 1334 CATAGCTTGTTGGTCCATCCAGAGTA 1308

Db 22 HisSerLeuValProSerArgVal 30
|||||
RESULT 6
ID MUTS_RICCN STANDARD; PRT; 890 AA.
AC 0921I9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA mismatch repair protein muts.
GN MUTS OR RCO401.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxId=781;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Malish 7;
RC MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audie S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissensbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
CC -1- FUNCTION: This protein is involved in the repair of mismatches in
CC DNA. It is possible that it carries out the mismatch recognition
CC step. This protein has a weak ATPase activity (By similarity).
CC -1- SIMILARITY: Belongs to the DNA mismatch repair muts family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008604; AL02939.1; -;
DR PIR; A97750; A97750.
DR HAMAP; MF_00096; -, 1.
DR InterPro; IPR005748; MutS1.
DR InterPro; IPR000432; MutS_C.
DR InterPro; IPR007860; MutS_I.
DR InterPro; IPR007896; MutS_III.
DR InterPro; IPR007861; MutS_IV.
DR InterPro; IPR007695; MutS_N.
DR Pfam; PF01624; MutS_I; 1.
DR Pfam; PF05188; MutS_II; 1.
DR Pfam; PF05192; MutS_III; 1.
DR Pfam; PF05190; MutS_IV; 1.
DR Pfam; PF00488; MutS_V; 1.
DR ProDom; PD001263; MutS_C; 1.
DR SMART; SM00534; MUTSAC; 1.
DR SMART; SM00533; MUTSD; 1.
DR TIGRfam; TIGR01070; mutS1; 1.
DR PROSITE; PS00486; DNA_MISMATCH_REPAIR_2; 1.
KW DNA repair; ATP-binding; DNA-binding; Complete proteome.
FT NP BIND 645 652
SQ SEQUENCE 890 AA; 100376 MW; 5016F26921D16792 CRC64;

Alignment Scores:
Pred. No.: 2.5 Length: 890
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: 1 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x MUTS_RICCN (1-890)
QY 1406 ATGCTGCTTCTCTTATCAGACTG 1380
|||||

Db 572 MetLeuAlaSerSerLeuSerGlyLeu 580

RESULT 7

HEM3_EUGGR

ID HEM3_EUGGR STANDARD; PRT; 480 AA.

AC P13446; 1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Porphobilinogen deaminase, chloroplast precursor (EC 2.5.1.61) (PBG)

OS (Hydroxymethylbilane synthase) (HMB) (Pre-urophorphyrinogen synthase).

OC Euglena gracilis.

OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.

OX NCBI_TaxID=3039;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN=Z / UTEX 753.

RX MEDLINE=90005485; PubMed=2477247;

RA Sharif A.U., Smith A.G., Abell C.;

RT "Isolation and characterization of a cDNA clone for a chlorophyll

RT synthesis enzyme from Euglena gracilis. The chloroplast enzyme

RT hydroxymethylbilane synthase (porphobilinogen deaminase) is

RT synthesised with a very long transit peptide in Euglena.";

RL Eur. J. Biochem. 164:353-359(1989).

CC -1- FUNCTION: Tetrapolymerization of the monopyrrole PBG into the

CC hydroxymethylbilane precursorporphyrinogen in several discrete steps.

CC -1- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =

CC hydroxymethylbilane + 4 NH(3).

CC -1- COFACTOR: Covalently binds a dipyrromethane cofactor to which the

CC porphobilinogen subunits are added.

CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; fourth step.

CC Involved in chlorophyll biosynthesis.

CC -1- SIMILARITY: Belongs to the HMBs family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X15743; CA33759.1; -.

DR PIR: S06109; IBBG.

DR HSSP: P06983; IPDA.

DR InterPro: IPR000860; Porphobil deam.

DR Pfam: PF01379; Porphobil deamC.1.

DR Pfam: PF03900; Porphobil deamC.1.

DR PRINTS: PR00151; PORPHBDMNASE.

DR ProDom: PD002745; Porphobil deam; 1.

DR TIGRfam: TIGR00212; hemc; 2.

DR PROSITE: PS00533; PORPHOBILINOGEN DEAM; 1.

KW Porphyrin biosynthesis; Chlorophyll biosynthesis; Transferase;

KW Chloroplast; Transit peptide.

FT TRANSIT 1 139 CHLOROPLAST.

FT CHAIN 140 480 PORPHOBILINOGEN DEAMINASE.

FT BINDING 395 395 PYRROMETHANE COFACTOR (BY SIMILARITY).

SEQUENCE 480 AA; 51743 MW; 269CE6CC195C0F3A CRC64;

Alignment Scores:

Pred. No.: 29
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.20%
DB: 1
Matches: 480
Conservative: 8
Mismatch: 0
Indels: 0
Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x HEM3_EUGGR (1-480)

QY 1722 CTACGAGGAGGAGGAGGAGGAGTACTC 1745

DB 279 LeuArgArgGlnCysGlnLeuLeu 286

RESULT 8

AGP4_MOUSE

ID AGP4_MOUSE STANDARD; PRT; 509 AA.

AC Q9VWH6;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Angiotensin-4 precursor (ANG-4) (ANG-3).

GN ANGPT4 OR AGP4 OR ANG3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Myoblasts, and uterus;

RX MEDLINE=99162530; PubMed=10051567;

RA Valenzuela D.M., Griffiths J.A., Rojas J., Aldrich T.H., Jones P.F.,

RA Zhou H., McClain J., Copeland N.G., Gilbert D.J., Jenkins N.A.,

RA Huang T., Papadopoulos N., Malsompiere P.C., Davis S.,

RT Yancopoulos G.D.;

RT "Angiotensins 3 and 4: diverging gene counterparts in mice and

RT humans.";

RL Proc. Natl. Acad. Sci. U.S.A. 96:1904-1909(1999).

CC -1- FUNCTION: BINDS TO TYROSINE-PROTEIN KINASE RECEPTOR TIE2 AND MAY

CC ACT AS AN ANTAGONIST.

CC -1- SUBCELLULAR LOCATION: Secreted (Probable).

CC -1- TISSUE SPECIFICITY: Widely expressed.

CC -1- SIMILARITY: Contains 1 fibrinogen C-terminal domain.

CC -1- CAUTION: Was originally (Ref.1) called angiotensin-3.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF113707; AAD21586.1; -.

DR HSSP: P02671; IEPD.

DR MGI: MGI:136887; Agp14.

DR InterPro: IPR002181; Fibrinogen_C.

DR Pfam: PF00147; fibrinogen_C; 1.

DR SMART: SM00186; FBG; 1.

DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; 1.

KW Coiled coil; Signal.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 509 ANGIOPOIETIN-4.

FT DOMAIN 181 269 COILED COIL (POTENTIAL).

FT DOMAIN 294 480 FIBRINOGEN-LIKE.

FT DISULFD 297 326 BY SIMILARITY.

FT DISULFD 450 463 BY SIMILARITY.

SEQUENCE 509 AA; 57805 MW; 9B5A74A20A664F6 CRC64;

Alignment Scores:

Pred. No.: 28.8
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.20%
DB: 1
Matches: 509
Conservative: 8
Mismatch: 0
Indels: 0
Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AGP4_MOUSE (1-509)

QY 1083 CGTGCCAGCGTGCCAGCGTGCA 1106

DB 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 9

VPS_BTIV0

ID VPS_BTIV0

AC P07389;

STANDARD;

PRT;

526 AA.

DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Outer capsid protein VP5.
GN 56 OR M5.
OS Bluesongue virus (serotype 10 / isolate USA).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
NCBI_TaxID=10900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=6198636; PubMed=3009695;
RA Purdy M.A., Ritter G.D., Roy P.;
RT "Nucleotide sequence of cDNA clones encoding the outer capsid
RT protein, VP5, of bluesongue virus serotype 10.";
RT J. Gen. Virol. 67:957-962(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90345726; PubMed=2166648;
RA Roy P., Marshall J.J.A., French T.J.;
RT "Structure of the bluesongue virus genome and its encoded proteins.";
RT Curr. Top. Microbiol. Immunol. 162:43-87(1990).
CC -1- FUNCTION: THE VP5 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP2)
CC WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID.
CC -1- SIMILARITY: Belongs to the reoviruses VP5 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL; D12532; BAA02095.1; -.
DR PIR; A25419; PSXR10.
DR InterPro; IPR000145; Orbl_VP5.
DR Pfam; PF00901; Orbl_VP5; 1.
KM Coat protein.
SQ SEQUENCE 526 AA; 59162 MW; 29AE556BDB55522C CRC64;
SO
Alignment Scores:
Pred. No.: 28.7 Length: 526
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 1 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x VP5_BT10 (1-526)
QY 98 CTATCAATTGCGACACTTTTTC 75
DB 435 Leuserleuserleuserleuser 442
RESULT 10
VGLF_P1IHC STANDARD; PRT; 555 AA.
AC P12605;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Human parainfluenza 1 virus (strain C39).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Respirovirus.
NCBI_TaxID=11210;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89045674; PubMed=2847427;
RA Merzon J.R., Hull R.A., Bates M.K., Kaseel J.A.;
RT "Molecular cloning and sequence determination of the fusion protein

RT gene of human parainfluenza virus type 1.";
RL Virology 167:97-105(1988).
CC -1- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL; M22347; AAA46800.1; -.
DR PIR; A31287; VGNZ11.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KM Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT FT 1 21
FT CHAIN 22 555
FT CHAIN 22 112
FT CHAIN 113 555
FT DISULFID 66 195
FT TRANSMEM 113 138
FT TRANSMEM 498 519
FT CARBOHYD 100 100
FT CARBOHYD 241 241
FT CARBOHYD 529 529
FT CARBOHYD 552 552
SQ SEQUENCE 555 AA; 60785 MW; 6F0DF62C969B9F0 CRC64;
SO
Alignment Scores:
Pred. No.: 28.5 Length: 555
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 1 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x VGLF_P1IHC (1-555)
QY 1193 AGCTTATGCTGCTTCAGCCTG 1170
DB 13 Serleuleuserleuserleu 20
RESULT 11
GOX_ASPNG STANDARD; PRT; 605 AA.
ID GOX_ASPNG
AC P13006;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glucose oxidase precursor (EC 1.1.3.4) (Glucose oxyhydrase) (GOD)
DE (Beta-D-glucose:oxygen 1-oxido-reductase).
GN GOX.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=NRRL 3;
RX MEDLINE=90005974; PubMed=2792372;
RA Kriechbaum M., Heilmann H.J., Wientjes F.J., Hahn M., Jany K.D.,
RA Gassen H.G., Sharif F., Alaedddinoglu G.;
RT "Cloning and DNA sequence analysis of the glucose oxidase gene from
RT Aspergillus niger NRRL-3.";
RL FEBS Lett. 255:63-66(1989).
RN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE=90154060; PubMed=2406261;
RA Frederick K.R., Tung J., Emerick R.S., Mastarz F.R., Chamberlain S.H.,
RA Vasavada A., Rosenberg S., Chakraborty S., Schopfer L.M., Massey V.,
RT "Glucose oxidase from *Aspergillus niger*. Cloning, gene sequence,
RT secretion from *Saccharomyces cerevisiae* and kinetic analysis of a
RT yeast-derived enzyme.";
RL J. Biol. Chem. 265:3793-3802(1990).
RN [3]
RP SEQUENCE OF 1-32 FROM N.A.
RX MEDLINE=91168301; PubMed=2076553;
RA Whittington H., Kerry-Williams S., Bidgood K., Dodsworth N.,
RA Peberdy J., Dobson M., Hinchliffe E., Ballance D.J.,
RT "Expression of the *Aspergillus niger* glucose oxidase gene in *A.*
RT *niger*, *A. nidulans* and *Saccharomyces cerevisiae*.";
RL Curr. Genet. 18:531-536(1990).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=9133782; PubMed=8421298;
RA Hecht H.-J., Kalisz H.M., Hendle J., Schmid R.D., Schomburg D.,
RT "Crystal structure of glucose oxidase from *Aspergillus niger* refined
RT at 2.3-A resolution.";
RL J. Mol. Biol. 229:153-172(1993).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=99234348; PubMed=10216293;
RA Wohlfahrt G., Wilt S., Hendle J., Schomburg D., Kalisz H.M.,
RA Hecht H.-J.,
RT "1.8 and 1.9-A resolution structures of the Penicillium amagaaklense
RT and *Aspergillus niger* glucose oxidases as a basis for modelling
RT substrate complexes.";
RL Acta Crystallogr. D 55:969-977(1999).
CC -1- CATALYTIC ACTIVITY: Beta-D-glucose + O(2) = D-glucono-1,5-lactone
CC + H(2)O(2).
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THIS ENZYME IS WIDELY APPLIED FOR THE DETERMINATION
CC OF GLUCOSE IN BODY FLUIDS AND IN REMOVING RESIDUAL GLUCOSE OR
CC OXYGEN FROM FOODS AND BEVERAGES. FURTHERMORE, GLUCOSE
CC OXIDASE-PRODUCING MOUTHS SUCH AS *ASPERGILLUS* AND *PENICILLIUM*
CC SPECIES ARE USED FOR THE BIOLOGICAL PRODUCTION OF GLUCONIC ACID.
CC -1- SIMILARITY: Belongs to the GMC oxidoreductase family.
CC -1- DATABASE: NAME=Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/cgp/".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X16061; CAA34197.1; -;
CC EMBL: J05242; AAA32695.1; -;
CC EMBL: X56443; CAA39826.1; -;
CC PIR: A35459; A35459.
CC PDB: 1GAL; 31-OCT-93.
CC PDB: 1CR3; 26-MAR-99.
CC InterPro: IPR000172; GMC_oxred.
CC InterPro: IPR007867; GMC_oxred_C.
CC Pfam: PF05199; GMC_oxred_C; 1.
CC Pfam: PF00732; GMC_oxred_N; 1.
CC PROSITE: PS00623; GMC_OXRED_1; 1.
CC PROSITE: PS00624; GMC_OXRED_2; 1.
CC Oxidoreductase; Flavoprotein; FAD; Glycoprotein; Signal; 3D-structure.
CC FT SIGNAL 1 22
CC FT CHAIN 1 22 OR 48.
CC FT NP_BIND 23 605 GLUCOSE OXIDASE.
CC FT ACT_SITE 543 543 FAD (ADP PART) (PROBABLE).
CC FT DISULFID 186 228 POTENTIAL.
CC FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .).
CC FT

FT HELIX 26 29
FT STRAND 31 31
FT HELIX 34 36
FT TURN 37 39
FT STRAND 41 47
FT HELIX 51 60
FT TURN 61 62
FT TURN 64 65
FT STRAND 68 72
FT TURN 78 79
FT HELIX 82 85
FT STRAND 86 86
FT HELIX 87 89
FT TURN 90 95
FT TURN 97 98
FT STRAND 99 99
FT STRAND 102 103
FT TURN 108 110
FT STRAND 115 116
FT STRAND 119 119
FT TURN 122 123
FT HELIX 124 127
FT TURN 128 128
FT STRAND 132 132
FT HELIX 138 146
FT TURN 147 148
FT TURN 151 152
FT HELIX 155 165
FT STRAND 166 168
FT HELIX 173 178
FT TURN 179 179
FT HELIX 184 186
FT STRAND 193 196
FT STRAND 204 204
FT TURN 206 206
FT HELIX 207 216
FT TURN 217 219
FT STRAND 222 222
FT TURN 226 227
FT STRAND 233 235
FT STRAND 239 239
FT STRAND 241 241
FT TURN 243 244
FT STRAND 246 247
FT HELIX 250 254
FT TURN 255 255
FT HELIX 256 258
FT TURN 259 260
FT TURN 262 263
FT STRAND 264 268
FT STRAND 271 278
FT STRAND 285 294
FT TURN 295 296
FT STRAND 297 309
FT TURN 313 315
FT HELIX 316 322
FT TURN 323 324
FT STRAND 325 325
FT STRAND 328 331
FT HELIX 332 334
FT STRAND 335 335
FT STRAND 340 340
FT TURN 344 345
FT STRAND 348 349
FT STRAND 352 360
FT HELIX 362 364
FT STRAND 367 367
FT STRAND 369 375
FT HELIX 376 380
FT HELIX 381 383
FT HELIX 384 393
FT HELIX 395 404
FT TURN 405 406

FT HELIX 411 426
FT TURN 427 428
FT STRAND 431 438
CC 440 441
FT TURN 442 449
FT STRAND 456 460
FT HELIX 465 467
FT STRAND 470 473
FT TURN 476 477
FT HELIX 480 497
FT TURN 498 498
FT HELIX 500 504
FT TURN 505 505
FT STRAND 506 511
FT HELIX 513 515
FT TURN 518 519
FT HELIX 522 528
FT STRAND 529 531
FT STRAND 534 535
FT STRAND 540 540
FT TURN 542 543
FT STRAND 545 545
FT HELIX 548 550
FT TURN 551 551
FT STRAND 554 554
FT TURN 555 557
FT STRAND 559 560
FT TURN 561 562
FT STRAND 564 564
FT STRAND 566 568
FT TURN 571 572
FT HELIX 583 603

Alignment Scores:

Pred. No.: 28.2 Length: 605
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x GOX_ASFNG (1-605)

QY 386 AGTCTGCTGTTCCCTGCGCC 409

Db 7 SerSerLeuValValSerLeuAla 14

RESULT 12

UL35_HCMVA STANDARD; PRT; 640 AA.

AC P16766;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Hypochemical protein UL35.
GN UL35.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90269039; PubMed=2161319;
RA Chee W.S., Bankier A.T., Beck S., Brown C.M., Cerny R.,
RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
RA Peddie B., Satchwell S.C., Tomlinson P., Weston K.M., Barrett B.G.,
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -1- SIMILARITY: BELONGS TO THE UL35 FAMILY.
CC -1- SIMILARITY: BELONGS TO THE UL35 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X17403; CAA35394.1; -
DR PIR; S09798; Q08EU2.
DR InterPro; IPR006731; Herpes_pp85.
DR Pfam; PF04637; Herpes_pp85; 1.
KW Hypochemical protein.
FT CARBOHYD 208
SQ SEQUENCE 640 AA; 72529 MW; 8FCCFA3C9F631C1C CRC64;

Alignment Scores:

Pred. No.: 28.1 Length: 640
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x UL35_HCMVA (1-640)

QY 1190 TTATTGCTGCTTCAGCCGCTCT 1167

Db 547 LeuLeuLeuSerSerLeuSer 554

RESULT 13

HUTH_CAEEL STANDARD; PRT; 677 AA.

AC P20502;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable histidine ammonia-lyase (EC 4.3.1.3) (Histidase).
GN F47810.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RA Harris B.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: L-histidine = urocanate + NH(3).
CC -1- PATHWAY: Histidine degradation; first step.
CC -1- PTM: Contains an active site 4-methylidene-imidazole-5-one (MIO),
CC which is formed autocatalytically by cyclization and dehydration
CC of residues Cys-Ser-Gly (By similarity).
CC -1- SIMILARITY: Belongs to the PAL / histidase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z68004; CAA31982.1; -
DR PIR; T22333; T22333.
DR HSSP; P21310; 1B8F.
DR WormPep; F47810.2; CE03352.
DR InterPro; IPR005921; Huth.
DR InterPro; IPR008948; L-Aspartase-like.
DR InterPro; IPR001106; Phe/His_NH3lyase.
DR Pfam; PF00221; PAL; 1.
DR TIGRFAMs; TIGR01225; huth; 1.
DR PROSITE; PS00488; PAL_HISTIDASE; 1.
KW Lyase; Histidine metabolism.
FT CROSSLINK 269 271 5-imidazolone (Cys-Gly)

```

FT      MOD_RES      270      270      (By similarity).
FT      SEQUENCE      677 AA; 74634 MW; B64CC5B097AAC4F CRC64;
SQ
Alignment Scores:
Pred. No.:      27.9      Length:      677
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      1.20%      Indels:      0
DB:              1      Gaps:      0
US-10-029-345A-108_COPY_532_2532 (1-2001) x HUTH_CAEEL (1-677)
OY      137 AACGATTAATATCAAGCTCCCA 160
DB      564 LysPtoleilleSerThalapro 591
RESULT 14
VIRL_AGRU      STANDARD;      PRT;      835 AA.
AC      P07167;
DT      01-APR-1998 (Rel. 07, Created)
DT      01-APR-1998 (Rel. 07, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Limited host range virA protein (EC 2.7.3.-) (LHR virA).
GN      VIRA.
OS      Agrobacterium tumefaciens.
OC      Plasmid pTiA6.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX      NCBI_TaxID=358;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      ST0AIN=A348, and A856;
RX      MEDLINE=87246518; PubMed=3595559.
RA      Leroux B., Yanozsky M.F., Winans S.C., Ward J.E., Ziegler S.F.,
RA      Neeter E.W.;
RT      "Characterization of the virA locus of Agrobacterium tumefaciens: a
RT      transcriptional regulator and host range determinant.";
RL      EMO J. 6:849-856(1987).
CC      -1- SIMILARITY: Contains 1 histidine kinase domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X05241; CAA28868.1; -.
DR      PIR; A27211; A27211.
DR      InterPro; IPR003594; ATPbind_Artpase.
DR      InterPro; IPR004358; Bact_sens_pr_C.
DR      InterPro; IPR003661; His_KinA_N.
DR      InterPro; IPR005467; His_Kinase.
DR      InterPro; IPR001789; Response_reg.
DR      Pfam; PF02518; HATPase_c; 1.
DR      Pfam; PF00512; HisKA_1.
DR      PRINTS; PR00344; BCTRLSENSOR.
DR      SMART; SM00387; HATPase_C; 1.
DR      SMART; SM00388; HisKA_1.
DR      SMART; SM00448; REC; 1.
DR      PROSITE; PS50109; HIS_KIN_1.
KW      Sensory transduction; Transferase; Kinase; Phosphorylation;
KW      Plasmid; Crown gall tumor; Transmembrane.
FT      TRANSMEM      19      37
FT      DOMAIN      476      698      HISTIDINE KINASE.
FT      MOD_RES      479      479      PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ      SEQUENCE      835 AA; 92443 MW; 93ADBA1E2B5F801 CRC64;

```

```

Alignment Scores:
Pred. No.:      27.2      Length:      835
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      1.24%      Indels:      0
DB:              1      Gaps:      0
US-10-029-345A-108_COPY_532_2532 (1-2001) x VIRL_AGRU (1-835)
OY      545 ATGAGCTCGTGTGAGCAGCATCT 522
DB      148 lIeSerSerLeuNarThirser 155
RESULT 15
RPA1_RAT
ID      RPA1_RAT      STANDARD;      PRT;      1716 AA.
AC      O54889;
DT      15-DEC-1998 (Rel. 37, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      DNA-directed RNA polymerase I largest subunit (EC 2.7.7.6) (RNA
DE      polymerase I 194 kDa subunit) (RPA194).
GN      POLR1A OR RPO1-4 OR RPA1.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=98086309; PubMed=9422795;
RX      Hannan R.D., Hempel W.M., Cavanaugh A., Arino T., Dimitrov S.I.,
RA      Moses T., Kochdum L.;
RT      "Affinity purification of mammalian RNA polymerase I. Identification
RT      of an associated kinase.";
RL      J. Biol. Chem. 273:1257-1267(1998).
CC      -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC      of DNA into RNA using the four ribonucleoside triphosphates as
CC      substrates. RNA polymerase I is essentially used to transcribe
CC      ribosomal DNA units.
CC      -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC      {RNA} (N).
CC      -1- SUBUNIT: Each class of RNA polymerase is assembled from 9 to 14
CC      different polypeptides. This subunit is the largest component of
CC      RNA polymerase I.
CC      -1- SUBCELLULAR LOCATION: Nuclear.
CC      -1- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
CC      found in eukaryotic nuclei: polymerase I for the ribosomal RNA
CC      precursor, polymerase II for the mRNA precursor, and polymerase
CC      III for 5S and tRNA genes.
CC      -1- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF025425; AAB94601.1; -.
DR      PIR; T14103; T14103.
DR      InterPro; IPR000722; RNA_pol_A.
DR      InterPro; IPR007080; RNA_pol_Rpb1_1.
DR      InterPro; IPR007066; RNA_pol_Rpb1_3.
DR      InterPro; IPR007083; RNA_pol_Rpb1_4.
DR      InterPro; IPR007081; RNA_pol_Rpb1_5.
DR      InterPro; IPR006592; RNA_polA_N.
DR      Pfam; PF04997; RNA_pol_Rpb1_1; 1.
DR      Pfam; PF00623; RNA_pol_Rpb1_2; 1.
DR      Pfam; PF04983; RNA_pol_Rpb1_3; 1.
DR      Pfam; PF05000; RNA_pol_Rpb1_4; 1.
DR      Pfam; PF04998; RNA_pol_Rpb1_5; 1.

```


DR SMART: SM00663; RPO1A N.1.
 KM Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
 KW Zinc-finger; Nuclear protein. C2H2-TYPE (POTENTIAL).
 FT ZN_FING 64 80
 SQ SEQUENCE 1716 AA; 194191 MW; EB8E15BC33E60941 CRC64;
 Alignment Scores:
 Pred. No.: 25 Length: 1716
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 1 Gaps: 0
 US-10-029-345a-108_COPY_532_2532 (1-2001) x RPA1_RAT (1-1716)
 QY 386 AGTTCTCTCGTTGTTTCCTGCGC 409
 DB 1537 SerSerLeuValValSerLeuAla 1544
 RESULT 16
 RPA1_HUMAN
 ID RPA1_HUMAN STANDARD; PRT; 1717 AA.
 AC 095602; Q9UEH0; Q9UEF7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE DNA-directed RNA polymerase I largest subunit (EC 2.7.7.6) (RNA
 DE polymerase I 194 kDa subunit) (RPA194).
 GN POLR1A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_TaxID=9606;
 RX NCBL_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA Wang D., Stetler D.A.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 SEQUENCE OF 1623-1717 FROM N.A.
 RC TISSUE=Uterus;
 RA Koehler K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
 of DNA into RNA using the four ribonucleoside triphosphates as
 substrates. RNA polymerase I is essentially used to transcribe
 ribosomal DNA units.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 (RNA) (N).
 CC -1- SUBUNIT: Each class of RNA polymerase is assembled from 9 to 14
 different polypeptides. This subunit is the largest component of
 RNA polymerase I.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
 found in eukaryotic nuclei: polymerase I for the ribosomal RNA
 precursor, polymerase II for the mRNA precursor, and polymerase
 III for 5S and tRNA genes.
 CC -1- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U33460; AAC98959.1; -;
 CC EMBL: U33823; AAD09356.1; -;
 CC EMBL: U33824; AAD09356.1; JOINED.
 CC EMBL: AL117467; CAB55942.1; -;
 CC PIR: T17252; T17252.
 CC SWISS-2DPAGE; 095602; HUMAN.

DR Genew; HGNC:17264; POLR1A.
 DR GO: 0005736; C:DNA-directed RNA polymerase I complex; NAS.
 DR InterPro: IPR000722; RNA_pol_A.
 DR InterPro: IPR007080; RNA_pol_Rpb1_1.
 DR InterPro: IPR007066; RNA_pol_Rpb1_3.
 DR InterPro: IPR007083; RNA_pol_Rpb1_4.
 DR InterPro: IPR007081; RNA_pol_Rpb1_5.
 DR InterPro: IPR006592; RNA_pol_N.
 DR Pfam: PF04997; RNA_pol_Rpb1_1; 1.
 DR Pfam: PF00623; RNA_pol_Rpb1_2; 1.
 DR Pfam: PF04983; RNA_pol_Rpb1_3; 1.
 DR Pfam: PF05000; RNA_pol_Rpb1_4; 1.
 DR Pfam: PF04998; RNA_pol_Rpb1_5; 1.
 DR SMART: SM00663; RPO1A N.1.
 KM Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
 KW Zinc-finger; Nuclear protein. C2H2-TYPE (POTENTIAL).
 FT ZN_FING 64 80
 FT CONFLICT 986 986 G -> A (IN REF. 1; AAD09356).
 FT CONFLICT 1372 1374 RAT -> TRI (IN REF. 1; AAD09356).
 SQ SEQUENCE 1717 AA; 194190 MW; 1143247B211F9BDB CRC64;
 Alignment Scores:
 Pred. No.: 25 Length: 1717
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 1 Gaps: 0
 US-10-029-345a-108_COPY_532_2532 (1-2001) x RPA1_HUMAN (1-1717)
 QY 386 AGTTCTCTCGTTGTTTCCTGCGC 409
 DB 1539 SerSerLeuValValSerLeuAla 1546
 RESULT 17
 RPA1_MOUSE
 ID RPA1_MOUSE STANDARD; PRT; 1717 AA.
 AC 035134;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA-directed RNA polymerase I largest subunit (EC 2.7.7.6) (RNA
 DE polymerase I 194 kDa subunit) (RPA194).
 GN POLR1A OR RPO1-4 OR RPA1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RX NCBL_TaxID=10090;
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97379976; PubMed=9236775;
 RA Seither P., Coy J.F., Pouška A., Grunnt I.;
 RL Molecular cloning and characterization of the cDNA encoding the
 RL largest subunit of mouse RNA polymerase I.;
 RL Mol. Gen. Genet. 255:180-186(1997).
 CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
 of DNA into RNA using the four ribonucleoside triphosphates as
 substrates. RNA polymerase I is essentially used to transcribe
 ribosomal DNA units.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 (RNA) (N).
 CC -1- SUBUNIT: Each class of RNA polymerase is assembled from 9 to 14
 different polypeptides. This subunit is the largest component of
 RNA polymerase I.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
 found in eukaryotic nuclei: polymerase I for the ribosomal RNA
 precursor, polymerase II for the mRNA precursor, and polymerase
 III for 5S and tRNA genes.
 CC -1- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; AF000938; AAB6718.1; -
 DR PIR; T13961; T13961.
 DR MGD; MG1:1096397; Rpol-4.
 DR InterPro; IPR000722; RNA_pol_A.
 DR InterPro; IPR007080; RNA_pol_Rpb1_1.
 DR InterPro; IPR007086; RNA_pol_Rpb1_3.
 DR InterPro; IPR007083; RNA_pol_Rpb1_4.
 DR InterPro; IPR007081; RNA_pol_Rpb1_5.
 DR InterPro; IPR006592; RNA_pol_N.
 DR Pfam; PF04997; RNA_pol_Rpb1_1.
 DR Pfam; PF00623; RNA_pol_Rpb1_2.
 DR Pfam; PF04983; RNA_pol_Rpb1_3.
 DR Pfam; PF05000; RNA_pol_Rpb1_4.
 DR Pfam; PF04998; RNA_pol_Rpb1_5.
 DR SMART; SM00663; RPOLA_N.1.
 DR Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
 KW Zinc-finger; Nuclear protein. CH2-TYPE (POTENTIAL).
 FT ZN FING
 SQ SEQUENCE 1717 AA; 194079 MW; 77C3F23F27FD29BE CRC64;

Alignment Scores:
 Pred. No.: 25 Length: 1717
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x RPA1_MOUSE (1-1717)

OY 386 AGTCTCGCTGCTTCCCTGCGCC 409
 DB 1538 SerSerLeuValValSerLeuAla 1545

RESULT 18
 POLN_SINDV STANDARD; PRT; 2512 AA.
 ID POLN_SINDV STANDARD; PRT; 2512 AA.
 AC P03317;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nonstructural polypeptide (P270) [Contains: Nonstructural protein
 DE NSP1; Nonstructural NSP2; Nonstructural NSP3; Nonstructural NSP4].
 OS Sindbis virus (strain HRSF).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus.
 OC NCBI_TaxID=11034;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84148439; PubMed=6322438;
 RA Straus E.G., Rice C.M., Straus J.H.;
 RT "Complete nucleotide sequence of the genomic RNA of Sindbis virus";
 RL Virology 133:92-110(1984).
 RN [2]
 RP SEQUENCE OF 1-54 FROM N.A.
 RX MEDLINE=83268700; PubMed=6308269;
 RA Ou J.H., Straus E.G., Straus J.H.;
 RT "The 5'-terminal sequences of the genomic RNAs of several
 RL alphaviruses".
 RL J. Mol. Biol. 168:1-15(1983).
 RN [3]
 RP SEQUENCE OF 1429-2512 FROM N.A.
 RX MEDLINE=83299955; PubMed=6577423;
 RA Straus E.G., Rice C.M., Straus J.H.;
 RT "Sequence coding for the alphavirus nonstructural proteins is
 RT interrupted by an opal termination codon.";

RL Proc. Natl. Acad. Sci. U.S.A. 80:5271-5275(1983).
 RN [4]
 RP SEQUENCE OF 2431-2512 FROM N.A.
 RX MEDLINE=83039346; PubMed=6291034;
 RA Ou J.H., Rice C.M., Dalgarno L., Straus E.G., Straus J.H.;
 RT "Sequence studies of several alphavirus genomic RNAs in the region
 RT containing the start of the subgenomic RNA".
 RL Proc. Natl. Acad. Sci. U.S.A. 79:5235-5239(1982).
 CC -1- FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UGA OCCURS
 CC BETWEEN THE CODONS FOR 1896-TYR AND 1897-LEU.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; J02363; AAA96975.1; ALT_SEQ.
 DR PIR; A03917; MNWVS.
 DR MEROPS; C09.001; -
 DR InterPro; IPR002589; A1pD.
 DR InterPro; IPR002620; Peptidase_C9.
 DR InterPro; IPR001788; RNA_dep_RNAPol2.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR006066; Viral_helicase1.
 DR Pfam; PF01661; A1pD; 1.
 DR Pfam; PF01707; Peptidase_C9; 1.
 DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
 DR Pfam; PF01443; Viral_helicase1; 1.
 DR SMART; SM00506; A1pD; 1.
 DR PolyProtein; Nonstructural protein; RNA-binding; Helicase.
 FT CHAIN 1 540 NONSTRUCTURAL PROTEIN NSP1.
 FT CHAIN 541 1347 NONSTRUCTURAL PROTEIN NSP2.
 FT CHAIN 1348 1896 NONSTRUCTURAL PROTEIN NSP3.
 FT CHAIN 1897 2512 NONSTRUCTURAL PROTEIN NSP4.
 SQ SEQUENCE 2512 AA; 279546 MW; F3656FC8BBA95726 CRC64;

Alignment Scores:
 Pred. No.: 23.9 Length: 2512
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x POLN_SINDV (1-2512)

OY 1550 CCAGCTTCCTTTGCGCCTTTCCA 1573
 DB 445 ProAlaSerPheSerAlaPhePro 452

RESULT 19
 POLN_SINDV STANDARD; PRT; 2514 AA.
 ID POLN_SINDV STANDARD; PRT; 2514 AA.
 AC P27283;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nonstructural polypeptide (P270) [Contains: Nonstructural protein
 DE NSP1; Nonstructural NSP2; Nonstructural NSP3; Nonstructural NSP4].
 OS Sindbis virus (subtype Ockelbo / strain Edabyn 82-5).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus.
 OC NCBI_TaxID=31699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220725; PubMed=1673813;
 RA Shirako Y., Niklasson B., Dalrymple J.M., Straus E.G., Straus J.H.;

```

RL Mamm. Genome 6:746-748(1995).
[2]
RP MUTAGENESIS.
RC STRAIN=Sprague-Dawley, and Fischer 344/N;
RX MEDLINE=95148647; PubMed=7846077;
RA Katsuhiko H., Watanabe M., Ushijima T., Toyota M., Imai K.,
RA Weisburger J. H., Sugimura T., Nagao M.;
RT "Specific 5'-GGGA-3'-->5'-GGA-3' mutation of the Apc gene in rat colon
RT tumors induced by 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:910-914(1995).
CC -1- FUNCTION: Tumor suppressor. Promotes rapid degradation of CTNNB1
CC and participates in Wnt signaling. APC activity is correlated with
CC its phosphorylation state (By similarity).
CC -1- SUBUNIT: Forms homooligomers. Associates with catenins. Binds
CC axin (By similarity).
CC -1- PTM: Phosphorylated by GSK3B (By similarity).
CC -1- SIMILARITY: Contains 7 ARM repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; D38629; BAA07609.1; -.
CC HSSP; Q02248; 3BCT.
CC InterPro; IPR008938; ARM.
CC InterPro; IPR000225; Armadillo.
CC Pfam; PF00514; Armadillo_seg; 4.
CC SMART; SM00185; ARM; 5.
CC PROSITE; PS50176; ARM_REPEAT; 1.
CC Wnt signaling pathway; Anti-oncogene; Phosphorylation; Coiled coil;
KW Repeat.
FT DOMAIN 1 728 LEU-RICH.
FT DOMAIN 1 62 COILED COIL (POTENTIAL).
FT DOMAIN 125 260 COILED COIL (POTENTIAL).
FT REPEAT 451 493 ARM 1.
FT REPEAT 503 545 ARM 2.
FT REPEAT 546 589 ARM 3.
FT REPEAT 590 636 ARM 4.
FT REPEAT 637 681 ARM 5.
FT REPEAT 682 723 ARM 6.
FT REPEAT 724 765 ARM 7.
FT DOMAIN 739 2831 SER-RICH.
FT DOMAIN 1130 1155 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 1556 1575 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 1864 1891 HIGHLY CHARGED.
FT MUTAGEN C->R: IN AN IQ-INDUCED COLON TUMOR.
SQ SEQUENCE 2842 AA; 310530 MW; 3CB82BA8A34EBF47 CRC64;
DB:

Alignment Scores:
Pred. No.: 23.6 Length: 2842
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x APC_RAT (1-2842)
QY 959 TTCTCAGGTGACAGCAGTTGAGT 936
DB 1268 Pheserargybserserleuser 1275

RESULT 21
APC_HUMAN STANDARD; PRT; 2843 AA.
ID APC_HUMAN
AC P25054; O15162; O15163;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)

```

DE Adenomatous polyposis coli protein (APC protein).
GN APC OR DP2.5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91335210; PubMed=1651562;
RA Kinzler K.W., Nilbert M.C., Su L.-K., Vogelstein B., Bryan T.M.,
RA Levy D.B., Smith K.J., Preisinger A.C., Hedge P., McKechnie D.,
RA Flinniar R., Markham A., Groffen J., Boguski M.S., Altschul S.F.,
RA Horii A., Ando H., Miyoshi Y., Miki Y., Nishisho I., Nakamura Y.,
RT "Identification of FAP locus genes from chromosome 5q21.";
RL Science 253:661-665 (1991).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Fetal brain;
RX MEDLINE=91330307; PubMed=1678319;
RA Joslyn G., Carlson M., Thliveris A., Albertsen H., Gelbert L.,
RA Samowitz W., Groden J., Stevens J., Spirio L., Robertson M.,
RA Sargent L., Kravcho K., Wolff E., Burt R., Hughes J.P.,
RA Warrington J., McPherson J.D., Wasmuth J., Le Fashier D.,
RA Abderrahim H., Cohen D., Leppert M., White R.,
RT "Identification of deletion mutations and three new genes at the
RL familial polyposis locus.";
RL Cell 66:601-613 (1991).
RN [3]
RP ASSOCIATION WITH CATENINS.
RX MEDLINE=94082295; PubMed=8259519;
RA Su L.-K., Vogelstein B., Kinzler K.W.,
RT "Association of the APC tumor suppressor protein with catenins.";
RL Science 262:1734-1737 (1993).
RN [4]
RP DISEASE.
RX MEDLINE=95174843; PubMed=7661930;
RA Hamilton S.R., Liu B., Parsons R.E., Papadopoulos N., Jen J.,
RA Powell S.M., Krush A.J., Berk T., Cohen Z., Iacu B., Burger P.C.,
RA Wood P.A., Tagi F., Booker S.V., Petersen G.M., Offerhaus G.J.A.,
RA Teremette A.C., Giardiello F.M., Vogelstein B., Kinzler K.W.,
RT "The molecular basis of Turcot's syndrome.";
RL New Engl. J. Med. 332:839-847 (1995).
RN [5]
RP DISEASE.
RX MEDLINE=97094176; PubMed=8940264;
RA Eccles D.M., van der Luijt R.B., Breukel C., Bullman H., Bunyan D.,
RA Fisher A., Barber J., du Boulay C., Primrose J., Burn J., Fodde R.,
RT "Hereditary desmoid disease due to a frameshift mutation at codon 1924
RL of the APC gene.";
RL Am. J. Hum. Genet. 59:1193-1201 (1996).
RN [6]
RP DISEASE.
RX MEDLINE=20243021; PubMed=10782927;
RA Couture J., Mitri A., Lagace R., Smits R., Berk T., Bouchard H.-L.,
RA Fodde R., Altman B., Bapat B.,
RT "A germline mutation at the extreme 3'-prime end of the APC gene
RT results in a severe desmoid phenotype and is associated with
RL overexpression of beta-catenin in the desmoid tumor.";
RL Clin. Genet. 57:205-212 (2000).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 2-55.
RX MEDLINE=20384842; PubMed=10926498;
RA Day C.L., Alber T.,
RT "Crystal structure of the amino-terminal coiled-coil domain of the
RL APC tumor suppressor.";
RL J. Mol. Biol. 301:147-156 (2000).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 1021-1035 IN COMPLEX WITH
RP CTNNB1.
RX MEDLINE=21564054; PubMed=11707392;
RA Ekiof Spink K., Fridman S.G., Weis W.I.,
RT "Molecular mechanisms of beta-catenin recognition by adenomatous
RT polyposis coli revealed by the structure of an APC-beta-catenin
RT complex.";
RT EMBO J. 20:6203-6212 (2001).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.90 ANGSTROMS) OF 2034-2049 IN COMPLEX WITH
RP AXIN.
RX MEDLINE=20271867; PubMed=10811618;
RA Spink K.E., Polakis P., Weis W.I.,
RT "Structural basis of the axin-adenomatous polyposis coli
RL interaction.";
RL EMBO J. 19:2270-2279 (2000).
RN [10]
RP REVIEW ON VARIANTS.
RX MEDLINE=94154728; PubMed=8111410;
RA Nagase H., Nakamura Y.,
RT "Mutations of the APC (adenomatous polyposis coli) gene.";
RL Hum. Mutat. 2:425-434 (1993).
RN [11]
RP VARIANTS FAP.
RX MEDLINE=91335211; PubMed=1651563;
RA Nishisho I., Nakamura Y., Miyoshi Y., Miki Y., Ando H., Horii A.,
RA Koyama K., Utsunomiya J., Baba S., Hedge P., Markham A., Krush A.J.,
RA Petersen G.M., Hamilton S.R., Nilbert M.C., Levy D.B., Bryan T.M.,
RA Preisinger A.C., Smith K.J., Su L.-K., Kinzler K.W., Vogelstein B.,
RT "Mutations of chromosome 5q21 genes in FAP and colorectal cancer
RL patients.";
RL Science 253:665-669 (1991).
RN [12]
RP VARIANTS FAP.
RX MEDLINE=93265030; PubMed=1338904;
RA Miyoshi Y., Nagase H., Ando H., Ichii S., Nakatsuru S., Aoki T.,
RA Miki Y., Mori T., Nakamura Y.,
RT "Somatic mutations of the APC gene in colorectal tumors: mutation
RL cluster region in the APC gene.";
RL Hum. Mol. Genet. 1:229-233 (1992).
RN [13]
RP VARIANTS FAP.
RX MEDLINE=93244793; PubMed=1338691;
RA Nakatsuru S., Yanagisawa A., Ichii S., Tahara E., Kato Y.,
RA Nakamura Y., Horii A.,
RT "Somatic mutation of the APC gene in gastric cancer: frequent
RT mutations in very well differentiated adenocarcinoma and signet-ring
RL cell carcinoma.";
RL Hum. Mol. Genet. 1:559-563 (1992).
RN [14]
RP VARIANTS FAP TRP-1348, AND VARIANTS ASP-1118; MET-1292; VAL-1304 AND
RP SER-2502.
RX MEDLINE=93250848; PubMed=1338764;
RA Nagase H., Miyoshi Y., Horii A., Aoki T., Petersen G.M.,
RA Vogelstein B., Maher E., Ogawa M., Maruyama M., Utsunomiya J.,
RA Baba S., Nakamura Y.,
RT "Screening for germ-line mutations in familial adenomatous polyposis
RT patients: 61 new patients and a summary of 150 unrelated patients.";
RL Hum. Mutat. 1:467-473 (1992).
RN [15]
RP VARIANTS FAP TRP-99.
RC TISSUE=Peripheral blood lymphocytes;
RX MEDLINE=95134544; PubMed=7833149;
RA Dobbie Z., Spycher M., Huertliman R., Ammann R., Ammann T., Roth J.,
RA Mueller A., Mueller H., Scott R.J.,
RT "Mutational analysis of the first 14 exons of the adenomatous
RT polyposis coli (APC) gene.";
RL Eur. J. Cancer 30A:1709-1713 (1994).
RN [16]
RP VARIANTS FAP GUY-722.
RX MEDLINE=95135430; PubMed=7833931;
RA Stella A., Montero M., Reeta N., Marchese C., Susca F., Gentile M.,
RA Romio L., Pilla S., Prete F., Marani C., Guanti G.,
RT "Four novel mutations of the APC (adenomatous polyposis coli) gene in
RT FAP patients.";
RL Hum. Mol. Genet. 3:1687-1688 (1994).
RN [17]
RP ERRATUM.
RA Stella A., Montero M., Reeta N., Marchese C., Susca F., Gentile M.,

RA Romio L., Pilla S., Prete F., Mareni C., Guanti G.;
 RL Hum. Mol. Genet. 3:1918-1918(1994).
 RN [18]
 RP VARIANT PAP ILE-171.
 RX MEDLINE=97144176; PubMed=8990002;
 RA van der Luit R.B., Meera Khan P., Vaseen H.F.A., Tops C.M.J.,
 van Leeuwen-Cornelisse I.S.J., Wijnen J.T., van der Klift H.M.,
 plug R.J., Griffioen G., Podde R.;
 RT "Molecular analysis of the APC gene in 105 Dutch kindreds with
 familial adenomatous polyposis: 67 germline mutations identified by
 DGGE, PRT, and southern analysis."
 RL Hum. Mutat. 9:7-16(1997).
 RN [19]
 RP VARIANTS COLORRECTAL CARCINOMA THR-880; ILE-890 AND VAL-1508.
 RX MEDLINE=98080146; PubMed=9419979;
 RA Miyaki M., Nishio J., Konishi M., Kikuchi-Yanoshta R., Tanaka K.,
 Murakawa M., Nagato M., Chong J.-M., Koike M., Terada T., Kawahara Y.,
 Fukuyama A., Tomiyama J., Chuganji Y., Momoi M., Urunomiya J.;
 RT "Drastic genetic instability of tumors and normal tissues in Turcot
 syndrome."
 RL Oncogene 15:2877-2881(1997).
 RN [20]
 RP VARIANT LYS-1307.
 RX MEDLINE=98400248; PubMed=9731522;
 RA Redetom M., Nathanson K.L., Yuan Z.Q., Neuhausen S.L., Saragopan J.,
 Wong N., Yang D., Nafa D., Abrahamson J., Ozcelik H.,
 Antin-Ozerkis D., Andrulis T., Daly M., Pinsky L., Schrag D.,
 Gallinger S., Kaback M., King M.-C., Woodage T., Brody L.C.,
 Godwin A., Warner E., Weber B., Foulkes W., Offit K.;
 RT "The APC I1307K allele and breast cancer risk."
 RL Nat. Genet. 20:13-14(1998).
 RN [21]
 RP VARIANTS LYS-1307 AND GLN-1317.
 RX MEDLINE=98393712; PubMed=9724771;
 RA Praying I.M., Beck N.E., Ilyas M., Dove-Edwin I., Goodman P.,
 Pack K., Bell J.A., Williams C.B., Hodgson S.V., Thomas H.J.W.,
 Talbot I.C., Bodmer W.F., Tomlinson I.P.M.;
 RT "The APC variants I1307K and E1317Q are associated with colorectal
 tumors, but not always with a family history."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:10722-10727(1998).
 RN [22]
 RP VARIANT LYS-1307.
 RX MEDLINE=98400259; PubMed=9731533;
 RA Woodgate T., King S.W., Wacholder S., Hartge P., Struwing J.P.,
 McAdams M., Iaken S.J., Tucker M.A., Brody L.C.;
 RT "The APC I1307K allele and cancer risk in a community-based study of
 Ashkenazi Jews."
 RL Nat. Genet. 20:62-65(1998).

Alignment Scores:
 Pred. No.: 23.6 Length: 2843
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.24% Indels: 0
 DB: 1 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x APC_HUMAN (1-2843)
 QY 959 TTCTCCAGTGCAGCAGCTTGACT 936
 DB 1271 PheSerArgCysSerSerLeuSer 1278

RESULT 22
 APC_MOUSE
 ID APC_MOUSE STANDARD; PRT; 2845 AA.
 AC 061315; 062044;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Adenomatous polyposis coli protein (APC protein) (mAPC).
 GN APC.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS.
 RX STRAIN=C57BL/6J, and CAST/EI; TISSUE=Brain;
 RA MEDLINE=92263101; PubMed=1350108;
 Su L.-K., Kinzler K.W., Vogelstein B., Preisinger A.C., Moser A.R.,
 Luongo C., Gould K.A., Dove W.F.;
 RT "Multiple intestinal neoplasia caused by a mutation in the murine
 homolog of the APC gene."
 RL Science 256:668-670(1992).
 RN [2]
 RP ERRATUM.
 RA Su L.-K., Kinzler K.W., Vogelstein B., Preisinger A.C., Moser A.R.,
 Luongo C., Gould K.A., Dove W.F.;
 RL Science 256:1114-1114(1992).
 RN [3]
 RP SEQUENCE OF 1-45 FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Liver;
 RA Dicker F., Lambertz S., Reilmair A., Ballhausen W.G.;
 RT "The murine APC gene: alternative splicing of 5' untranslated
 region segments."
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=94061824; PubMed=8242607;
 RA Oshima M., Sugiyama H., Kitagawa K., Taketo M.;
 RT "APC gene messenger RNA: novel isoforms that lack exon 7."
 RL Cancer Res. 53:5589-5591(1993).
 CC -1- FUNCTION: Tumor suppressor. Promotes rapid degradation of CTNNB1
 and participates in Wnt signaling. APC activity is correlated with
 its phosphorylation state (By similarity).
 CC -1- SUBUNIT: Forms homooligomers. Associates with catenine. Binds
 axin (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=4;
 Name=1;
 IsoId=Q61315-1; Sequence=Displayed;
 Name=2;
 IsoId=Q61315-2; Sequence=VSP_004116;
 Name=3;
 IsoId=Q61315-3; Sequence=VSP_004117;
 Name=4;
 IsoId=Q61315-4; Sequence=VSP_004116, VSP_004117;
 CC -1- TISSUE SPECIFICITY: Expressed in liver, spleen, kidney, heart,
 lung, brain, stomach, intestine, testis and ovary.
 CC -1- PTM: Phosphorylated by GSK3B (By similarity).
 CC -1- SIMILARITY: Contains 7 ARM repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M88127; AAB59632.1; -
 DR EMBL; U02937; AAA03443.1; -
 DR FIR; I49505; I49505.
 DR HSSP; Q02248; 3BCT.
 DR MGD; MGJ:88039; APC.
 DR GO; GO:0005737; Cytoplasm; IDA.
 DR GO; GO:0005634; Cytoplasm; IDA.
 DR GO; GO:0008013; F-beta-catenin binding; IDA.
 DR GO; GO:0009952; Anterior/posterior pattern formation; IMP.
 DR GO; GO:0009796; P-axis specification; IMP.
 DR GO; GO:0009953; P-dorsal/ventral pattern formation; IMP.
 DR GO; GO:0016055; P:Wnt receptor signaling pathway; IDA.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR000225; Armadillo.

DR Pfam: PF00514; Armadillo_seg; 4.
 DR SMART; SM00185; ARM; 5.
 DR PROSITE; PS00176; ARM_REPEAT; 1.
 KW mt signaling pathway; Anti-oncogene; Phosphorylation;
 KW Alternative splicing; Repeat; Coiled coil.
 FT DOMAIN 1 61
 FT DOMAIN 125 245
 FT DOMAIN 1 728
 FT REPEAT 451 493
 FT REPEAT 503 545
 FT REPEAT 546 589
 FT REPEAT 590 636
 FT REPEAT 637 681
 FT REPEAT 682 723
 FT REPEAT 724 765
 FT DOMAIN 739 2834
 FT DOMAIN 1130 1156
 FT DOMAIN 1556 1575
 FT DOMAIN 1864 1891
 FT VARSPIC 243 276
 FT VARSPIC 310 410
 FT VARSPIC 120 120
 FT VARIANT 493 493
 FT VARIANT 797 797
 FT VARIANT 1330 1330
 FT VARIANT 1618 1618
 FT VARIANT 2294 2294
 FT VARIANT 2496 2496
 FT VARIANT 2523 2523
 FT VARIANT 2813 2813
 SQ SEQUENCE 2845 AA; 311086 MW; 145CA73CF570A499 CRC64;

Alignment Scores:
 Pred. No.: 23.6
 Score: 8.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.24%
 DB: 1
 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x APC_MOUSE (1-2845)

QY 959 TTCTCAGGTGACGACCTTGAGT 936
 DB 1269 PheserargCysserSerIenSer 1276

RESULT 23
 ABGA_CLOLO STANDARD; PRT; 473 AA.
 AC 046130;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 6-phospho-beta-glucosidase (EC 3.2.1.86).
 GN ABGA.
 OS Clostridium longisporum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 NCBI_TaxID=1523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B6405;
 RA Brown G.D.; Thomson J.A.;
 RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: 6-phospho-beta-D-glucoside-(1,4)-D-glucose +
 H(2)O = D-glucose 6-phosphate + D-glucose.
 CC -1- SIMILARITY: Belongs to family 1 of glycosyl hydrolases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; L49336; AAC05714.1; -.
 DR HSSP; P11546; 1PBG.
 DR InterPro; IPR001360; Glyco_hydro_1.
 DR Pfam; PF00232; Glyco_hydro_1; 1.
 DR PRINTS; PR00131; GHYDRLASE1.
 DR ProDom; PD000650; Glyco_hydro_1; 1.
 DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 1.
 DR PROSITE; PS00653; GLYCOSYL_HYDROL_F1_2; 1.
 KW Hydrolase; Glycosidase.
 FT ACT SITE 174 174
 FT ACT SITE 366 366
 FT ACT SITE 366 366
 SQ SEQUENCE 473 AA; 54413 MW; 9862BA0023903ECD CRC64;

Alignment Scores:
 Pred. No.: 314
 Score: 7.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.03%
 DB: 1
 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x ABGA_CLOLO (1-473)

QY 1791 ATGACCTTGTCCTCCGCAAGT 1771
 DB 420 IleaspleuValserAlaser 426

RESULT 24
 ATPB_PTEES STANDARD; PRT; 473 AA.
 AC 003080;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE ATP synthase beta chain (EC 3.6.3.14) (Fragment).
 GN ATPB.
 OS Pteridium esculentum.
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Moniliformopses; Filicophyta; Filicopsida; Filicales;
 OC Dennstaedtiaceae; Pteridium.
 NCBI_TaxID=32102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wolf P.G.;
 RT "Evaluation of atpB nucleotide sequences for phylogenetic studies of
 RT ferns and other pteridophytes";
 RL Am. J. Bot. 84:1429-1440(1997).
 CC -1- FUNCTION: Produces ATP from ADP in the presence of a proton
 CC gradient across the membrane. The beta chain is the catalytic
 CC subunit.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 H(+) (Out).
 CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
 CC core - and CF(0) - the membrane proton channel. CF(1) has five
 CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
 CC has three main subunits: a, b and c.
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
 CC -1- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```
DR EMBL; U93834; AAB51742.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR005722; ATP_synthf_beta.
DR InterPro; IPR00793; ATPase_a/b_C.
DR InterPro; IPR00194; ATPase_a/b_centre.
DR InterPro; IPR004100; ATPase_a/bN.
DR InterPro; IPR009005; F1_ATPase_a/bN.
DR Pfam; PF00006; ATP-synt_ab; 1.
DR Pfam; PF00306; ATP-synt_ab_C; 1.
DR Pfam; PF02874; ATP-synt_ab_N; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRPFAMs; TIGR01039; atpD; 1.
DR PROSITE; PS00152; ATPase_ALPHA_BETA; 1.
DR ATP synthase; Chloroplast; Thylakoid; Membrane; CF(1);
DR Hydrolyase; ATP-binding; Hydrogen ion transport.
FT NON_TER 1
FT NP_BIND 172 179 ATP (POTENTIAL).
FT NON_TER 473
FT TER 473
SQ SEQUENCE 473 AA; 50434 MW; 5F305968B8B90029 CRC64;

Alignment Scores:
Pred. No.: 314 Length: 473
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.05% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x ATPB_PTEES (1-473)
OY 193 GTGTTATTACAGAGTCATC 213
DB 180 ValLeu11eHnGrLeu11e 186

RESULT 25
ASCB_ECOLI STANDARD; PRT; 474 AA.
ID _ASCB_ECOLI
AC P24240; P78104; Q59375;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phospho-beta-glucosidase ascb (EC 3.2.1.86).
GN ASCB OR B2716.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92334140; PubMed=1630307;
RA Hall B.G., Xu L.;
RT "Nucleotide sequence, function, activation, and evolution of the
RT cryptic asc operon of Escherichia coli K12.";
RL Mol. Biol. Evol. 9:688-706(1992).
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.O.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: CAN HYDROLYSE SALICIN, CELLOBIOSE, AND PROBABLY
CC ARABUTIN.
CC -1- CATALYTIC ACTIVITY: 6-phospho-beta-D-glucoside-(1,4)-D-glucose +
CC H(2)O = D-glucose 6-phosphate + D-glucose.
CC -1- SIMILARITY: Belongs to family 1 of glycosyl hydrolases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M73326; AAA16430.1; -.
DR EMBL; U29579; AAA69226.1; ALT_INIT.
DR EMBL; AE000355; AAC75758.1; -.
DR PIR; H65051; H65051.
DR HSSP; P11546; 1PBG.
DR EcoGene; EG10085; ascb.
DR InterPro; IPR001360; Glyco_hydro_1.
DR Pfam; PF00232; Glyco_hydro_1; 1.
DR PRINTS; PR00131; GLHYDRLASE1.
DR PRODOM; PD000650; Glyco_hydro_1; 1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 1.
DR PROSITE; PS00653; GLYCOSYL_HYDROL_F1_2; 1.
DR Hydrolyase; Glycosidase; Complete proteome.
FT ACT_SITE 160 180 PROTON DONOR (POTENTIAL).
FT ACT_SITE 372 372 NUCLEOPHILE (BY SIMILARITY).
FT CONFLICT 405 406 EA -> GT (IN REF. 1).
FT CONFLICT 428 428 S -> C (IN REF. 1).
FT CONFLICT 455 456 RK -> HR (IN REF. 1).
SQ SEQUENCE 474 AA; 53935 MW; 02ACE6BBF211011 CRC64;

Alignment Scores:
Pred. No.: 314 Length: 474
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x ASCB_ECOLI (1-474)
OY 1791 ATGACTGTGTCGCGCAGT 1771
DB 422 l1eapbuevalseralaser 428

RESULT 26
PERI_MOUSE STANDARD; PRT; 475 AA.
ID _PERI_MOUSE
AC P15311; Q35688; Q35689;
DT 01-APR-1990 (Rel. 14, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Peripherin.
GN PRPH OR PRPH1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=liver;
RX MEDLINE=93192779; PubMed=1294287;
RA Karlov V., Landon F., Djabali K., Gros F., Portier M.M.;
RT "Structure of the mouse gene encoding peripherin: a neuronal
RT intermediate filament protein.";
RL Biol. Cell 76:43-48(1992).
RN 12
RP SEQUENCE FROM N.A. (ISOFORM SG).
RC STRAIN=FVB/N; TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheafer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang Y., Hsieh F.,
RA Diatchenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Umed T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
```

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maira M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 70-475 FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE=69356645; PubMed=2767051;
 RA Landon F., Lemontier M., Benarous R., Huc C., Fiszman M., Gros F.,
 RA Portier M.M.,
 RT "Multiple mRNAs encode peripherin, a neuronal intermediate filament
 RT protein".
 RL EMO J. 8:1719-1726(1989).
 CC -1- FUNCTION: Peripherin are class-III neuronal intermediate filament
 CC proteins.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=5g;
 CC IsoId=PI5331-1; Sequence=Displayed;
 CC Name=3u;
 CC IsoId=PI5331-2; Sequence=VSP_002466;
 CC Name=5b;
 CC IsoId=PI5331-3; Sequence=VSP_002467;
 CC -1- SIMILARITY: Belongs to the intermediate filament family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X15475; CAA3502.1; -
 DR EMBL; BC046291; AAH46291.1; -
 DR EMBL; X59840; CAA42499.1; -
 DR EMBL; X59840; CAA42500.1; -
 DR EMBL; X59840; CAA42501.1; -
 DR PIR; S14887; S14887.
 DR MGI; MGI:97774; Prphl.
 DR InterPro: IPR006821; Filament_head.
 DR InterPro: IPR001664; IF.
 DR InterPro: IPR002957; Keratin_1.
 DR Pfam; PF00038; Filament_1.
 DR Pfam; PF04732; Filament_head_1.
 DR PRINTS; PRO1248; TYPEKERATIN.
 DR PROSITE; PS00226; IF, 1.
 KW Intermediate filament; Coiled coil; Phosphorylation;
 KW Alternative splicing.
 FT DOMAIN 1 103 HEAD.
 FT DOMAIN 104 409 ROD.
 FT DOMAIN 410 475 TAIL.
 FT DOMAIN 104 136 COIL_1A.
 FT DOMAIN 137 147 LINKER_1.
 FT DOMAIN 148 243 COIL_1B.
 FT DOMAIN 244 266 LINKER_2.
 FT DOMAIN 267 409 COIL_2.
 FT MOD_RES 475 475 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 475 475 K -> KVEHMGNPGGPRVGHGHWEMRCASPGLSATQ
 FT VARSPLIC 294 294 (in isoform 3u).
 FT VARSPLIC 294 294 (FTId=VSP_002466.
 FT VARSPLIC 294 294 KVTESQKEQSHSDDKSSHSY -> LLRQPEL (in
 FT VARSPLIC 294 294 isoform 5b).
 FT VARSPLIC 294 294 (FTId=VSP_002467.
 FT VARSPLIC 294 294 777300B08A312953 CRC64;
 SQ SEQUENCE 475 AA; 54267 MM; 777300B08A312953 CRC64;

Alignment Scores:
 Pred. No.: 314
 Score: 7.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.05%
 Gaps: 0
 US-10-029-345a-108_COPY_532_2532 (1-2001) x PERI_MOUSE (1-475)
 QY 758 CCGCCACATCGCTATCGCCT 778
 Db 30 ProProProSerLeuSerPro 36
 |||||
 RESULT 27
 ID SIML YEAST STANDARD; PRT; 475 AA.
 AC P40472;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE SIML protein precursor.
 GN SIML OR PB3 OR YII123W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=9713266; PubMed=9169870;
 RA Churcher C.M., Bowman S., Baddock K., Bankier A., Brown D.,
 RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
 RA Harris D.E., Horenell T., Hunt S., Jags R., Jones M., Lye G.,
 RA Moute S., Odeli C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
 RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.,
 RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX".
 RL Nature 387:84-87(1997).
 RN [2]
 RP IDENTIFICATION.
 RX MEDLINE=96155611; PubMed=8574583;
 RA Dahmann C., Diffley J.F.X., Nasmyth K.A.;
 RT "S-phase-promoting cyclin-dependent kinases prevent re-replication by
 RT inhibiting the transition of replication origins to a pre-replicative
 RT state".
 RL Curr. Biol. 5:1257-1269(1995).
 CC -1- FUNCTION: REQUIRED FOR THE MAINTENANCE OF THE CLAS KINASE
 CC ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE SUN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z46833; CAA86869.1; -
 DR PIR; S49865; S49866.
 DR Geronline; 139658; -
 DR SGD; S0001385; SIML.
 DR GO; GO:0009277; C:cell wall (sensu Fungi); IDA.
 DR GO; GO:0000226; P:microtubule cytoskeleton organization and b. .; IGI.
 DR InterPro: IPR005556; SUN.
 DR Pfam; PF03856; SUN, 1.
 KW Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 475 SIML PROTEIN.
 FT DOMAIN 58 112 ALA-RICH.
 FT DOMAIN 80 203 SER-RICH.
 FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 475 AA; 48070 MM; 8780F0BC8478612P CRC64;

Alignment Scores:

Pred. No.: 314 Length: 475
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 1 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x SIM1_YEAST (1-475)

QY 1058 CCTGCTGCTCTGAGGTAGCA 1038
 |||||

Db 57 ProAlaAlaSerGIuValAla 63

RESULT 28

LXC2_PHOLE STANDARD; PRT; 478 AA.

AC P29236; RT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DE Acyl-CoA reductase (EC 1.2.1.50).

GN LUXC.

OC Bacteria; Photobacterium; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Photobacterium.

OC NCBI_TaxID=658;

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 25521;

RA MEDLINE=92007870; PubMed=1915359;

lee C.Y., Seltzer R.B., Meighen E.A.;

"The lux genes of the luminous bacterial symbiont, Photobacterium

RT leiognathi, of the ponyfish, Nucleotide sequence, difference in gene

RT organization, and high expression in mutant Escherichia coli.";

RL Bur. J. Biochem. 201:161-167(1991).

CC -1- FUNCTION: LUXC IS THE FATTY ACID REDUCTASE ENZYME RESPONSIBLE FOR

CC SYNTHESIS OF THE ALDEHYDE SUBSTRATE FOR THE LUMINESCENT REACTION

CC CATALYZED BY LUCIFERASE.

CC -1- CATALYTIC ACTIVITY: A long-chain aldehyde + CoA + NADPH = a

CC long-chain acyl-CoA + NADPH.

CC -1- PATHWAY: Bioluminescent fatty acid reduction system; third step.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC EMBL; M63594; AAA25616.1; -

DR PIR; S17836.

DR InterPro; IPR008670; LuxC.

DR Pfam; PF05893; LuxC; 1.

KW Luminescence; Oxidoreductase; NADP.

SO SEQUENCE 478 AA; 53713 MW; 4716F699BF7FE4A CRC64;

Alignment Scores:

Pred. No.: 313 Length: 478
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.05% Indels: 0
 DB: 1 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x LXC2_PHOLE (1-478)

QY 355 TCTGTTACCTGCTTGACGT 375
 |||||

Db 145 SerValHisLeuValAlaGly 151

RESULT 29

TR2_HUMAN STANDARD; PRT; 483 AA.

ID TR2_HUMAN

AC P13056; RT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last annotation update)

GN NR2C1 OR TR2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88339993; PubMed=3421977;

RA Chang C., Kokontis J.;

RT "Identification of a new member of the steroid receptor super-family

RT by cloning and sequence analysis.";

RL Biochem. Biophys. Res. Commun. 155:971-977(1988).

CC -1- FUNCTION: Orphan nuclear receptor.

CC -1- SUBCELLULAR LOCATION: Nuclear (potential).

CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR2

CC subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC EMBL; M21985; AAA3650.1; -

DR PIR; A31521; A31521.

DR HSSP; P19793; 2NLL.

DR TRANSFAC; T01662; -

DR Genew; HGNC:7971; NR2C1.

DR MIM; 601529; -

DR GO; GO:0003677; F:DNA binding; TAS.

DR GO; GO:0004872; F:receptor activity; TAS.

DR GO; GO:0003707; F:steroid hormone receptor activity; TAS.

DR InterPro; IPR000536; Hormone_rec_1lg.

DR InterPro; IPR001723; Sterhmr_receptor.

DR InterPro; IPR008946; Str_ncl_receptor.

DR InterPro; IPR001628; Znf_C4steroid.

DR Pfam; PF00104; hormone_rec; 1.

DR Pfam; PF00105; zfc4; 1.

DR PRINTS; PR00398; STRDHOMONER.

DR PRINTS; PR00047; STROIDPRNGER.

DR ProDom; PD000035; Znf_C4steroid; 1.

DR SMART; SM00399; Znf_C4; 1.

DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.

KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;

KW zinc-finger.

FT DNA_BIND 113 178 NUCLEAR_RECEPTOR-TYPE.

FT ZN_FING 113 133 C4-TYPE.

FT ZN_FING 149 173 C4-TYPE.

SO SEQUENCE 483 AA; 53042 MW; 293E21E317E05814 CRC64;

Alignment Scores:

Pred. No.: 313 Length: 483
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 1 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x TR2_HUMAN (1-483)

QY 487 GGGTGGCCCAATGTTGCA 467
 |||||

Db 437 GlyLeuAlaGlnCysTyrIpsln 443

```
RESULT 30
YML8_STRAW
ID YML8_STRAW STANDARD; PRT; 487 AA.
AC 082KZ4;
DT 10-OCT-2003 (Rel. 42, Created)
DE 10-OCT-2003 (Rel. 42, Last sequence update)
DN 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0276 protein SAV2218.
GN SAV2218.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Kikuchi H., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis; deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
CC -1- SIMILARITY: In the N-terminal section; belongs to the UPF0276
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP005029; BAC6929.1; -
DR HAMAP; MF 00697; fused. 1; -
DR InterPro; IPR007801; DUF692.
DR Pfam; PF05114; DUF692; 1.
KW Hypothetical protein; Complete proteome.
FT DOMAIN 1 285 UNKNOWN.
FT FT 286 487
SQ SEQUENCE 487 AA; 51637 MW; 0652ADB870BF369 CRC64;

Alignment Scores:
Pred. No.: 313
Score: 7.00 Length: 487
Percent Similarity: 100.00% Matches: 7
Best Local Similarity: 100.00% Conservative: 0
Query Match: 1.09% Mismatches: 0
DB: Indels: 0
Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x YML8_STRAW (1-487)
QY 1103 ACGCTGGGACGCTGGGACG 1083
DB 12 ThrLeuGlyThrLeuGlyThr 18

RESULT 31
VGLY_MOPEI
ID VGLY_MOPEI STANDARD; PRT; 489 AA.
AC P19240;
DT 01-NOV-1990 (Rel. 16, Created)
```

```
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glycoprotein polypeptide precursor [Contains: Glycoprotein G1;
DE Glycoprotein G2].
GN GPC.
OS Mopeia virus.
OC Viruses; ssRNA negative-strand viruses; Arenaviridae; Arenavirus.
OX NCBI_TaxID=11629;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=800150;
RX MEDLINE=91111973; PubMed=1989384;
RA Wilson S.M., Clegg J.C.S.;
RT "Sequence analysis of the S RNA of the African arenavirus Mopeia: an
RT unusual secondary structure feature in the intergenic region.";
RL Virology 180:543-552(1991).
CC -1- SIMILARITY: Belongs to the arenaviruses GPC protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M33879; AAC08700.1; -
DR PIR; A38546; VGXPMV.
DR InterPro; IPR01535; Arena glycoprot.
DR Pfam; PF00798; Arena glycoprot; 1.
KW Polypeptide; Glycoprotein; Envelope protein.
FT CHAIN 1 255 GLYCOPROTEIN G1.
FT FT 256 489
FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 108 108 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 222 222 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 363 363 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 489 AA; 55987 MW; B188B8327E170B1 CRC64;

Alignment Scores:
Pred. No.: 312
Score: 7.00 Length: 489
Percent Similarity: 100.00% Matches: 7
Best Local Similarity: 100.00% Conservative: 0
Query Match: 1.05% Mismatches: 0
DB: Indels: 0
Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x VGLY_MOPEI (1-489)
QY 971 CTGTCCTGCTGTCTCAGAGG 991
DB 240 LeuSerLeuLeuSerGlnArg 246

RESULT 32
YI10_YEAST
ID YI10_YEAST STANDARD; PRT; 491 AA.
AC P40499;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 56.3 Kda protein in SGAI-KTR7 intergenic region.
GN YI1090W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RX MEDLINE=97313266; PubMed=9169870;
RA Chicher C.M., Bowman S., Baddock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Horsnell T., Hunt S., Jagals K., Jones M., Lye G.,
RA Moul S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;
RL "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IX";
RM Nature 387:84-87(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: SOME, TO S.POMBE SPAC23C11.01.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 246728; CAA86704.1; -.
DR PIR; S49790; S49790.
DR Germonline; 139625; -.
DR SGD; S0001352; YII090W.
KM Hypothetical protein; Transmembrane.
FT TRANSMEM 14 34 POTENTIAL.
FT TRANSMEM 41 61 POTENTIAL.
FT TRANSMEM 75 95 POTENTIAL.
FT TRANSMEM 150 170 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 219 239 POTENTIAL.
FT TRANSMEM 295 315 POTENTIAL.
FT TRANSMEM 374 394 POTENTIAL.
FT TRANSMEM 414 434 POTENTIAL.
FT TRANSMEM 462 482 POTENTIAL.
SQ SEQUENCE 491 AA; 56332 MW; BAE655534E04239 CRC64;

Alignment Scores:
Pred. No.: 312 Length: 491
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x YII0_YEAST (1-491)
QY 83 ACTTTTCGTCCTTCCTTC 63
DB 353 ThpHeservatProleuser 359

RESULT 33
FLIC_SALTY STANDARD; PRT; 494 AA.
AC P06179; P97160; Q02871; Q56088;
DT 01-JAN-1988 (Rel. 06, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Flagellin (Phase-1-I flagellin).
GN FLIC OR FLAP OR HMG OR STM1959.
OS *Salmonella typhimurium*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; *Salmonella*.
OC NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=6059460; PubMed=2999134;
RA Jovy T.M.;
RL "The covalent structure of the phase-1 flagellar filament protein of
RT *Salmonella typhimurium* and its comparison with other flagellins";
RJ J. Biol. Chem. 260:15758-15761(1985).

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93139045; PubMed=8423149;
RA Okazaki N., Matsuo S., Saito K., Tomimaga A., Enomoto M.;
RT "Conversion of the *Salmonella* phase 1 flagellin gene *fliC* to the
RL phase 2 gene *fliB* on the *Escherichia coli* K-12 chromosome.";
RJ J. Bacteriol. 175:758-766(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RL "Complete genome sequence of *Salmonella enterica* serovar *Typhimurium*
RT LT2";
RM Nature 413:852-856(2001).
RN [4]
RP SEQUENCE OF 1-20 FROM N.A.
RX MEDLINE=90294298; PubMed=2193164;
RA Homma M., Derosier D.J., McNab R.M.;
RT "Flagellar hook and hook-associated proteins of *Salmonella*
RT typhimurium and their relationship to other axial components of the
RL flagellum";
RJ J. Mol. Biol. 213:819-832(1990).
RN [5]
RP SEQUENCE OF 1-19 FROM N.A.
RX MEDLINE=81054741; PubMed=6933466;
RA Zeig J., Simon M.;
RT "Analysis of the nucleotide sequence of an invertible controlling
RT element";
RJ Proc. Natl. Acad. Sci. U.S.A. 77:4196-4200(1980).
RN [6]
RP SEQUENCE OF 144-427 FROM N.A.
RX MEDLINE=90130292; PubMed=2404944;
RA Smith N.H., Seldner R.K.;
RT "Sequence invariance of the antigen-coding central region of the phase
RT 1 flagellar filament gene (*fliC*) among strains of *Salmonella*
RT typhimurium";
RJ J. Bacteriol. 172:603-609(1990).
RN [7]
RP SEQUENCE OF 475-494 FROM N.A.
RC STRAIN=LT2 / ATCC 23564;
RX MEDLINE=97311993; PubMed=9168604;
RA Burnens A.P., Stanley J., Sack R., Hunziker P., Brodard I.,
RA Nicolet J.;
RT "The flagellin N-methylase gene *fliB* and an adjacent serovar-specific
RT IS200 element in *Salmonella typhimurium*";
RJ Microbiology 143:1539-1547(1997).
RN [8]
RP FUNCTION: Flagellin is the subunit protein which polymerizes to
CC form the filaments of bacterial flagella.
CC -1- MISCELLANEOUS: Individual *salmonella* serotypes usually alternate
CC between the production of 2 antigenic forms of flagella, termed
CC phase-1 and phase-2, each specified by separate structural genes,
CC *fliC* and *fliB*.
CC -1- SIMILARITY: Belongs to the bacterial flagellin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M11332; AAA27072.1; -.
DR EMBL; D13689; BAA02846.1; -.
DR EMBL; AE008787; AAL20871.1; -.
DR EMBL; X51740; CAA6029.1; -.
DR EMBL; J01801; AAA27074.1; -.
DR EMBL; M33808; AAA27080.1; -.

EMBL; 254217; CAA09050.1; -.
 DR PIR; A24262; A24262.
 DR PIR; S16121; S16121.
 DR PDB; 1IO1; 15-APR-03.
 DR StyGene; SG10115; FLIC.
 DR InterPro; IPR001029; Flagellin C.
 DR InterPro; IPR001492; Flagellin N.
 DR Pfam; PF00700; Flagellin C.1.
 DR Pfam; PF00669; Flagellin N.1.
 DR PRINTS; PR00207; FLAGELLIN.
 DR Flagellum; Complete proteome, 3D-structure.
 FT INIT MET
 FT CONFLICT 0
 FT CONFLICT 126
 FT CONFLICT 132
 FT CONFLICT 132
 FT CONFLICT 214
 FT CONFLICT 260
 FT CONFLICT 276
 FT CONFLICT 283
 FT CONFLICT 337
 FT CONFLICT 337
 FT CONFLICT 345
 FT CONFLICT 353
 FT CONFLICT 374
 FT CONFLICT 381
 FT CONFLICT 386
 FT CONFLICT 389
 FT CONFLICT 393
 FT CONFLICT 399
 FT CONFLICT 402
 FT CONFLICT 447
 FT CONFLICT 447
 SQ SEQUENCE 494 AA; 51480 MW; 599EAD0DEA9DBEB8 CRC64;

Alignment Scores:

Pred. No.: 312
 Score: 7.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.09%
 DB: 1
 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x FLIC_SALTY (1-494)

QY 1743 GTAACGCACTGCTCGGTA 1723
 Db 300 ValThrglyThrAlaSerVal 306

RESULT 34

RP RNPL_HUMAN STANDARD; PRT; 494 AA.
 AC Q9HAI8; Q9HAI8; Q9HAI8; Q9HAI8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Arginyl aminopeptidase-like 1 (EC 3.4.11.-) (RNPEP-like protein).
 GN RNPEP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RX MEDLINE=20472315; PubMed=11017071;
 RA Horikawa Y., Oda N., Cox N.J., Li X., Orho-Melander M., Hara M.,
 RA Hinojosa Y., Lindner T.H., Mashima H., Schwarz P.E.H.,
 RA del Bosque-Platawa L., Horikawa Y., Oda Y., Yoshitachi I., Collila S.,
 RA Polonsky K.S., Wei S., Concannon P., Iwasaki N., Schlitz J.,
 RA Baler L.J., Bogardus C., Groop L., Boerwinkle E., Hais C.L.,
 RA Bell G.I.;
 RT "Genetic variation in the calpain 10 gene (CAPN10) is associated with
 RT type 2 diabetes";
 RL Nat. Genet. 26:163-175(2000).
 RP [2]
 RP SEQUENCE OF 53-494 FROM N.A.
 RP TISSUE=Pancreas;

EX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Hirschman R.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stadelton M., Soares M.B., Bonaldo M.F., Casavante T.L., Scheetz T.E.,
 RA Brownstein M.J., Urdin T.B., Toshiyuki S., Carninci P., Prange S.J.,
 RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bock S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Stamerich A., Schein J.E., Jones S.J.M., Skalska U., Smalins D.E.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [3]
 RP SEQUENCE OF 116-494 FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahara K., Masuho Y., Sasaki N.,
 RT "NEBO human cDNA sequencing project";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 242-494 FROM N.A.
 BC TISSUE=Lymph node;
 RA Bloecher H., Boecher M., Brandt P., Mewes H.-W., Weill B., Wiemann S.;
 CC Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -I- COPACTOR: Binds 1 zinc ion per subunit (by similarity).
 CC -I- TISSUE SPECIFICITY: Ubiquitous. Expressed at relatively higher
 CC levels in heart and skeletal muscle.
 CC -I- SIMILARITY: Belongs to peptidase family M1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF300795; AAG22080.1; -.
 DR EMBL; BC017301; AAH17301.1; ALT_INIT.
 DR EMBL; AK001668; BAA91823.1; ALT_INIT.
 DR EMBL; AL512754; CAC21674.1; -.
 DR Genew; HGNC:10079; RNPEP1.
 DR MIM; 605287; -.
 DR MEROPS; M01.022; -.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR001930; Peptidase_M1.
 DR Pfam; PF01433; Peptidase_M1.
 DR PRINTS; PR00756; BLADIPASE.
 DR PROSITE; PS00142; ZINC_PROTEASE; FALSE NEG.
 KW Aminopeptidase; Hydrolase; Zinc; Metalloprotease.
 FT METAL 122 122
 FT ACT SITE 123 123
 FT METAL 126 126
 FT METAL 145 145
 FT ACT SITE 211 211
 FT CONFLICT 378 378
 SQ SEQUENCE 494 AA; 55548 MW; 15872AD82E4E999D CRC64;

Alignment Scores:

Pred. No.: 312
 Score: 7.00
 Percent Similarity: 100.00%
 Length: 494
 Matches: 7
 Conservative: 0

Best Local Similarity: 100.00%
Query Match: 1.09%
DB: 1
Indels: 0
Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x RNPL_HUMAN (1-494)

QY 1445 TGGCTGTCTGAGGCTGCGC 1425
DB 142 TriptenSerCluicyneala 148

RESULT 35

ID HUTH_THEAC STANDARD; PRT; 496 AA.
AC 09H16;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Probable histidine ammonia-lyase (EC 4.3.1.3) (Histidase).

GN HUTH OR TA0242.

OS Thermoplasma acidophilum.

OC Archaea; Euryarchaeota; Thermoplasmatata; Thermoplasmatales;

OC Thermoplasmataceae; Thermoplasma.

OX NCBI_TaxID=2303;

RP SEQUENCE FROM N.A.

RC STRAIN=DSM 1728;

RX MEDLINE=20479972; PubMed=11029001;

RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,

RT Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumelster W.;

RT "The genome sequence of the thermophilic scavenger Thermoplasma

acidophilum.";

RL Nature 407:508-513(2000).

CC -1- CATALYTIC ACTIVITY: L-histidine = urocanate + NH(3).

CC -1- PATHWAY: Histidine degradation; first step.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Pocentia).

CC -1- PFM: Contains an active site 4-methylidazole-5-one (MIO),

CC which is formed autocatalytically by cyclization and dehydration

CC of residues Ala-Ser-Gly (By similarity).

CC -1- SIMILARITY: Belongs to the PAL / histidase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AL445063; CAC11387.1; -.

DR HSSP; P21310; 1B8F.

DR HAMAP; MF_00229; -; 1.

DR InterPro; IPR005921; Huth.

DR InterPro; IPR008948; L-Asparaginase-like.

DR InterPro; IPR001106; Ph/His_NH3lyase.

DR Pfam; PF00221; PAL; 1.

DR TIGRPFAM; TIGR01225; huth.1.

DR PROSITE; PS00488; PAL_HISTIDASE; FALSE_NEG.

DR Lyase; Histidine metabolism; Complete proceme.

KT CROSSLINK 141 143 5-Imidazoleone (Ala-Gly)

FT MOD_RES 142 142 (By similarity).

FT DHX (2,3-DIDEHYDROALANINE)

FT (BY SIMILARITY).

SEQUENCE 496 AA; 54148 MW; D5P0BF275477CA23 CRC64;

Alignment Scores:

Pred. No.: 312

Score: 7.00

Length: 496

Matches: 7

Conservative: 0

Best Local Similarity: 100.00%

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Mismatches: 0

Indels: 0

Query Match: 1.09%

Gaps: 0

DB: 1

US-10-029-345A-108_COPY_532_2532 (1-2001) x HUTH_THEAC (1-496)

QY 1383 ACTGGTTTCGGAGTCTGCTC 1363
DB 56 InrGlyPneGlySerLeuLeu 62

RESULT 36

ID MATK_LIOFO STANDARD; PRT; 503 AA.

AC 098370;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DE Maturase K (Intron maturase).

GN MATK.

OS Liquidambar formosana (Formosan gum).

OC Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Saxifragales; Hamamelidaceae; Liquidambar.

OX NCBI_TaxID=63359;

RN [1]

RP SEQUENCE FROM N.A.

RA Li J.-H., Bogle A.L., Klein A.S.;

RT "Phylogenetic inference in Altingioideae (Hamamelidaceae) based on

RT mark and its sequences."

RT Liquidambar (Hamamelidaceae)."

RT Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Probably assists in splicing chloroplast group II

CC introns (By similarity).

CC -1- SIMILARITY: Belongs to the intron maturase family 2. Matk

CC subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF015650; AAD10955.1; -.

DR EMBL; AF133221; AAG23349.1; -.

DR InterPro; IPR000442; Intron_mature2.

DR InterPro; IPR002866; Matk_N.

DR Pfam; PF01348; Intron_mature2; 1.

DR Pfam; PF01824; Matk_N; 1.

KW mRNA processing; Chloroplast.

SEQUENCE 503 AA; 59125 MW; C7C6A36D9B85CF35 CRC64;

Alignment Scores:

Pred. No.: 311

Score: 7.00

Length: 503

Matches: 7

Conservative: 0

Best Local Similarity: 100.00%

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Mismatches: 0

Indels: 0

Query Match: 1.05%

US-10-029-345A-108_COPY_532_2532 (1-2001) x MATK_LIOFO (1-503)

Alignment Scores:

Pred. No.: 1574

Score: 7.00

Length: 503

Matches: 7

Conservative: 0

Best Local Similarity: 100.00%

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Mismatches: 0

Indels: 0

Query Match: 1.09%

Gaps: 0

DB: 1

US-10-029-345A-108_COPY_532_2532 (1-2001) x MATK_LIOFO (1-503)

Alignment Scores:

Pred. No.: 1574

Score: 7.00

Length: 503

Matches: 7

Conservative: 0

Best Local Similarity: 100.00%

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Mismatches: 0

Indels: 0

Query Match: 1.09%

Gaps: 0

DB: 1

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
 NC NCB1_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Macon T.M., Bowman C.L., Barnstead M.E., Feldlyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
 RA Moffit K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
 RA Niernan W.C., White O., Eisen J.A., Salzberg S.L., Frazer C.M.,
 RA Venter J.C.,
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana."
 RL Nature 402:761-768(1999).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Karlín-Nemmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.W., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carrinai P.,
 RA Chao Q., Choy N., Enji A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashiki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamita A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamez R., Vaysberg W., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome."
 RL Science 302:842-846(2003).
 RN (3)
 RP SEQUENCE OF 4-507 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=99087420; PubMed=9872415;
 RA Kenzior A.L., Folk W.R.,
 RT "AtMS14 and Rbap48 WD-40 repeat proteins bind metal ions."
 RL FEBS Lett. 440:425-429(1998).
 RN (4)
 RP SEQUENCE OF 1-142 AND 452-507 FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=green alligues;
 RA Raynal M., Grellet F., Landie M., Meyer Y., Cooke R., Delseny M.,
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: Belongs to the WD-repeat RBAF6/RBAP48/MS11 family.
 CC -1- SIMILARITY: Contains 6 WD repeats.
 CC -1- CAUTION: Ref.4 (CAA85542) sequence differs from that shown due to
 CC a frameshift in position 136.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC
 CC EMBL; AC005917; AAD10151.2; ALT_INIT.
 CC EMBL; AY059799; AAL24281.1; -
 CC EMBL; AY057655; AAL15286.1; -
 CC EMBL; AY081447; AAM10009.1; -
 CC EMBL; AF028711; AAD03340.1; -
 CC EMBL; Z37286; CAA85542.1; ALT_FRAME.
 CC EMBL; Z37287; CAA85543.1; -
 CC PIR; G84577; G84577.
 CC InterPro: IPR001680; WD40.
 CC Pfam; PF00400; WD40; 5.

DR PRINTS; PR00320; GPROTEINRPT.
 DR Prodom; PD000018; WD40; 1.
 DR SMART; SM00320; WD40; 6.
 DR PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
 DR PROSITE; PS50082; WD_REPEATS_2; 3.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KM Repeat; WD repeat; Nuclear protein.
 FT REPEAT 162 193
 FT REPEAT 217 248
 FT REPEAT 290 321
 FT REPEAT 335 366
 FT REPEAT 384 415
 FT REPEAT 439 477
 FT DOMAIN 468 471
 FT CONFLECT 89 89
 FT CONFLECT 126 126
 FT CONFLECT 202 202
 FT CONFLECT 270 270
 FT CONFLECT 463 463
 FT CONFLECT 489 489
 FT CONFLECT 494 494
 SQ SEQUENCE 507 AA; 55759 MW; C37F8000F8B3397 CRC64;
 Alignment Scores:
 Pred. No.: 311
 Score: 7.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.05%
 DB: 1
 Gaps: 0
 US-10-029-345a-108_COPY_532_2532 (1-2001) x MS14_ARATH (1-507)
 QY 910 ACTGAGCATCAGCGCCAAAG 930
 Db 20 ThrGlyAlaSerGlyProlys 26
 RESULT 40
 PUR1_CHICK STANDARD; PRT; 510 AA.
 ID P28173;
 AC P28173;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 10-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Amino-phosphoribosyltransferase precursor (EC 2.4.2.14) (Glutamine
 DE dephosphoribosylpyrophosphate amidotransferase) (ATASE) (GPAT).
 GN GPAT OR GPAT.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NC NCB1_TaxID=9031;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91065928; PubMed=8336716;
 RA Zhou G., Dixon J.E., Zaikin H.,
 RT "Cloning and expression of avian glutamine
 RT phosphoribosylpyrophosphate amidotransferase. Conservation of a
 RT bacterial propylid sequence supports a role for posttranslational
 RT processing."
 RL J. Biol. Chem. 265:21152-21159(1990).
 RN (2)
 RP SEQUENCE OF 1-42 FROM N.A.
 RX MEDLINE=9330273; PubMed=8336716;
 RA Gavaldas A., Dixon J.E., Brayton K.A., Zaikin H.,
 RT "Coexpression of two closely linked avian genes for purine nucleotide
 RT synthesis from a bidirectional promoter."
 RL Mol. Cell. Biol. 13:4784-4792(1993).
 RN (3)
 RP ERRATUM.
 RA Gavaldas A., Dixon J.E., Brayton K.A., Zaikin H.,
 RL Mol. Cell. Biol. 13:7977-7977(1993).
 CC -1- CATALYTIC ACTIVITY: 5-phospho-beta-D-riboseylamine + diphosphate +

```

CC      L-glutamate = L-glutamine + 5-phospho-alpha-D-ribose 1-diphosphate
CC      + H(2)O.
CC      -1- COFACTOR: Binds 1 magnesium ion and 1 4Fe-4S cluster per subunit
CC      (By similarity).
CC      -1- PATHWAY: De novo purine biosynthesis, first step.
CC      -1- SUBUNIT: Homotetramer.
CC      -1- SIMILARITY: In the C-terminal section; belongs to the
CC      purine/pyrimidine phosphoribosyltransferase family.
CC      -1- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M6069; AAA62736.1; -.
CC      EMBL; L12533; AAA17895.1; -.
CC      PIR; A38337; A38337.
CC      HSSP; P00497; 1A00.
CC      MEROPS; C44.001; -.
CC      InterPro; IPR005854; Amd_phepho_trans.
CC      InterPro; IPR000583; GAtase_2.
CC      InterPro; IPR002375; Pr/PY_ip_transf.
CC      InterPro; IPR000836; PRTtransferase.
CC      Pfam; PR00310; GAtase_2; 1.
CC      Pfam; PR00156; Pribosyltran; 1.
CC      TIGRfam; TIGR01134; purf; 1.
CC      PROSITE; PS00403; PUR_PYR_PR_TRANSFER; 1.
CC      PROSITE; PS00443; GATASE_TYPE_II; 1.
CC      Purine biosynthesis; Transferase; Glycosyltransferase;
CC      Glutamine amidotransferase; Allosteric enzyme; Metal-binding;
CC      Magnesium; Iron-sulfur; 4Fe-4S.
CC      KEGG; K01111; PROBABLE.
CC      PROPEP; 1.
CC      FT CHAIN 12 510 AMIDOPHOSPHORIBOSYLTRANSFERASE.
CC      FT ACT SITE 12 12 GATASE (BY SIMILARITY).
CC      FT METAL 280 280 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC      FT METAL 327 327 MAGNESIUM (BY SIMILARITY).
CC      FT METAL 389 389 MAGNESIUM (BY SIMILARITY).
CC      FT METAL 390 390 MAGNESIUM (BY SIMILARITY).
CC      FT METAL 426 426 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC      FT METAL 496 496 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC      FT METAL 499 499 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC      FT METAL 499 499 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC      SQ SEQUENCE 510 AA; 56257 MW; F4371FB1FBC7C44 CRC64;

Alignment Scores:
Pred. No.: 311 Length: 510
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.05% Indels: 0
Gaps: 0
DB: 1

US-10-029-345A-108_COPY_532_2532 (1-2001) x PUR1_CHICK (1-510)
QY 1786 GCTCTATCTGCGCGGCGG 1806
DB 299 VAlTYrSeVaLrGArGArG 305

RESULT 41
HUTH_VIBCH STANDARD; PRT; 511 AA.
ID HUTH_VIBCH STANDARD; PRT; 511 AA.
AC Q9KSO4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Histidine ammonia-lyase (EC 4.3.1.3) (Histidase).
GN HUTH OR VC1202.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.

```

```

OX NCBI_TaxId=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RA MEDLINE=20406833; PubMed=10952501;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Ginn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bacs S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Frazer C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483 (2000).
CC -1- CATALYTIC ACTIVITY: L-histidine = urocanate + NH(3).
CC -1- PATHWAY: Histidine degradation; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- PTM: Contains an active site 4-methylidene-imidazole-5-one (MIO),
CC which is formed autocatalytically by cyclization and dehydration
CC of residues Ala-Ser-Gly (By similarity).
CC -1- SIMILARITY: Belongs to the PAL / histidase family.
CC -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AE004200; AAF94361.1; -.
CC      PIR; B82228; B82228.
CC      HSSP; P21310; 1B8F.
CC      TIGR; VC1202; -.
CC      DR HAMAP; MF_00229; -.
CC      DR InterPro; IPR005921; Hutch.
CC      DR InterPro; IPR008948; L-Aspartase-like.
CC      DR InterPro; IPR001066; Phe/His_NH3lyase.
CC      Pfam; PR00221; PAL; 1.
CC      DR TIGRfam; TIGR01225; huth; 1.
CC      DR PROSITE; PS00488; PAL_HISTIDASE; 1.
CC      KEGG; K01143; PAL; 1.
CC      FT Lyase; Histidine metabolism; Complete proteome.
CC      FT CROSSLINK 143 145 5-imidazolinone (Ala-Gly)
CC      FT MOD_RES 144 144 DHA (2,3-DIDEHYDROALANINE)
CC      FT MOD_RES 144 144 DHA (2,3-DIDEHYDROALANINE)
CC      SQ SEQUENCE 511 AA; 54816 MW; E872E0CB79FAF38A CRC64;

Alignment Scores:
Pred. No.: 311 Length: 511
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.05% Indels: 0
Gaps: 0
DB: 1

US-10-029-345A-108_COPY_532_2532 (1-2001) x HUTH_VIBCH (1-511)
QY 1853 AGAGCCCTTTGAAAGCAGT 1873
DB 453 ArGAlAPrOleuLYSeSer 459

RESULT 42
COXI_MYXGL STANDARD; PRT; 517 AA.
ID COXI_MYXGL STANDARD; PRT; 517 AA.
AC 021079;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cytochrome c oxidase polypeptide I (BC 1.9.3.1).
GN COXI OR COI.
OS Myxine glutinosa (Atlantic hagfish).

```


OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotretli; Myxiniiformes;
 OC Myxiniidae; Myxiniinae; Myxine.
 OX NCBI_TaxID=7769;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98210228; PubMed=9541532;
 RA Rasmussen A.S., Janke A., Arnason U.;
 RT "The mitochondrial DNA molecule of the hagfish (Myxine glutinosa) and
 vertebrate phylogeny.";
 RL J. Mol. Evol. 46:382-388(1998).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21534745; PubMed=11677623;
 RA Delabre C., Rasmussen A.S., Arnason U., Gachelin G.;
 RT "The complete mitochondrial genome of the hagfish Myxine glutinosa:
 unique features of the control region.";
 RL J. Mol. Evol. 53:634-641(2001).
 RN (3)
 RP SEQUENCE OF 1-8 FROM N.A.
 RX MEDLINE=97398704; PubMed=9254918;
 RA Delabre C., Barriel V., Tillet S., Janvier P., Gachelin G.;
 RT "The main features of the craniate mitochondrial DNA between the NDI
 and the COI genes were established in the common ancestor with the
 lancelet.";
 RL Mol. Biol. Evol. 14:807-813(1997).
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 chain that catalyzes the reduction of oxygen to water. Subunits
 1-3 form the functional core of the enzyme complex. CO I is the
 catalytic subunit of the enzyme. Electrons originating in
 cytochrome c are transferred via the copper A center of subunit 2
 and heme A of subunit 1 to the bimetallic center formed by heme A3
 and copper B.
 CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
 c + 2 H(2)O.
 CC -1- PATHWAY: Respiratory chain; terminal step.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 inner membrane. Contains 12 potential transmembrane domains.
 CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y15182; CAAT5481.1; -;
 DR EMBL: AJ404477; CAC20651.1; -;
 DR EMBL: Y09527; CAAT0718.1; -;
 DR PIR: T13818; T13818.
 DR HSSP: P00396; 20CC.
 DR InterPro: IPR000883; COX1.
 DR Pfam: PF00115; COX1; 1.
 DR PRINTS: PR01165; CYCOXIDASE1.
 DR PROSITE: PS00077; COX1; 1.
 KW Oxidoreductase; Copper; Mitochondrion; Transmembrane;
 KW Respiratory chain; Inner membrane.
 FT METAL 61 61 IRON (HEME A AXIAL LIGAND) (PROBABLE).
 FT METAL 240 240 COPPER B (PROBABLE).
 FT METAL 244 244 COPPER B (PROBABLE).
 FT METAL 290 290 COPPER B (PROBABLE).
 FT METAL 291 291 COPPER B (PROBABLE).
 FT METAL 376 376 IRON (HEME A3 AXIAL LIGAND) (PROBABLE).
 FT METAL 378 378 IRON (HEME A AXIAL LIGAND) (PROBABLE).
 FT CROSSLINK 240 244 1'-histidyl-3'-tyrosine (His-Tyr
 (By similarity)).
 SQ SEQUENCE 517 AA; 57947 MW; 286F869510D532AB CRC64;

Alignment Scores:

Pred. No.: 310 Length: 517
 Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 1 Gaps: 0
 US-10-029-345a-108_COPY_532_2532 (1-2001) x YHG9_YEAST (1-517)
 QY 1190 TTATTGCTGCTTCGAGCCTG 1170
 Db 111 LeuLeuLeuSerSerSerLeu 117
 RESULT 43
 YHG9_YEAST STANDARD; PRT; 518 AA.
 AC P38758;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 57.0 kDa protein in SOD2-RPL27A intergenic region.
 GN YHR009C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;
 OX NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 Kucaba T., Hillier L.W., Jier M., Johnston L., Langston Y.,
 Latreille P., Louis R.J., Macri C., Mardis E., Menges S., Mouser L.,
 Nhan M., Rifkin L., Riley L., St. Peter H., Trevaske E., Vaughan K.,
 Vaudin M., Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 VIII.";
 RL Science 265:2077-2082(1994).
 CC -1- SIMILARITY: TO S.POMBE SPAC1F5.03C AND SPAC610.06.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U10400; AAB68938.1; -;
 DR PIR: S45784; S45784.
 DR GerMOnline: J39326; -;
 DR SGD: S0001051; YHR009C.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 19 35 POTENTIAL.
 SQ SEQUENCE 518 AA; 57034 MW; 362C244F0F2C9959 CRC64;

Alignment Scores:

Pred. No.: 310 Length: 518
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.05% Indels: 0
 DB: 1 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x YHG9_YEAST (1-518)

QY 1190 AGCTCAAGCTCTCTCTCTC 1210

Db 179 SerSerSerValProSerLeu 185

RESULT 44
 YAV8_SCHPO STANDARD; PRT; 521 AA.
 ID YAV8_SCHPO

AC Q10177;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C27F1.08 in chromosome I.
 GN SPAC27F1.08.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 ON NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Vancura E., Aert R., Robben J., Gymnopoulos B.,
 RA Wellens J., Vancura E., Aert R., Robben J., Gymnopoulos B.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Gabor C., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambut R., Punnell B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McComb W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Uesery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880 (2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the NRAMP family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z69368; CA93297.1; -
 DR FJ: T38466; T38466.
 DR GeneDB_SpombE; SPAC27F1.08; -
 DR InterPro; IPR001046; Nrmamp.
 DR Pfam; PF01566; Nrmamp; 1.
 DR PRINTS; PR00447; NATRESASSCMP.
 DR ProDom; PD001861; Nrmamp; 1.
 DR TIGRfam; TIGR0197; nrmamp; 1.
 KW Hypothetical protein; Transmembrane; Transport.
 FT TRANSMEM 71 91 POTENTIAL.
 FT TRANSMEM 104 124 POTENTIAL.
 FT TRANSMEM 152 172 POTENTIAL.
 FT TRANSMEM 179 199 POTENTIAL.
 FT TRANSMEM 210 230 POTENTIAL.
 FT TRANSMEM 233 253 POTENTIAL.
 FT TRANSMEM 260 280 POTENTIAL.
 FT TRANSMEM 325 345 POTENTIAL.
 FT TRANSMEM 373 393 POTENTIAL.
 FT TRANSMEM 417 437 POTENTIAL.
 FT TRANSMEM 440 460 POTENTIAL.
 FT TRANSMEM 495 515 POTENTIAL.

SQ SEQUENCE 521 AA; 57641 MW; F14271E85FD7805A CRC64;
 Alignment Scores:
 Pred. No.: 310 Length: 521
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 1 Gaps: 0
 US-10-029-345a-108_COPY_532_2532 (1-2001) x YAV8_SCHPO (1-521)
 QY 903 CTTAATCTTCTCTCATAGTC 883
 Db 507 LeuAsnLeuLeuIleVal 513
 |||||
 RESULT 45
 ID HEMK_RICCN STANDARD; PRT; 524 AA.
 AC 092613;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Bifunctional methyltransferase [includes: Hemk protein homolog
 DE (BC 2.1.1.-) (M.RohemKP); tRNA (guanine-N(7)-)-methyltransferase
 DE (BC 2.1.1.33) (tRNA(m7G46)-methyltransferase)].
 GN HEMK OR RCI314.
 OS Rickettsia conorii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 ON NCBI_TaxId=781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Malish 7;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii";
 RL Science 293:2093-2098 (2001).
 CC -1- FUNCTION: Catalyzes the formation of N(7)-methylguanine at
 CC position 46 (m7G46) in tRNA (by similarity).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
 CC homocysteine + tRNA containing N(7)-methylguanine.
 CC -1- SIMILARITY: In the N-terminal section, belongs to the hemk family.
 CC -1- SIMILARITY: In the C-terminal section, belongs to the
 CC methyltransferase superfamily. Trmb family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE008678; AL03852.1; -
 DR HAMAP; MF_01057; fused; 1.
 DR InterPro; IPR004395; Cons_hypoth91.
 DR InterPro; IPR004556; HemK.
 DR InterPro; IPR003358; Methyltransf_4.
 DR InterPro; IPR002052; N6_Mtase.
 DR Pfam; PF02390; SAM bind.
 DR TIGRfam; TIGR00536; hemk_fam; 1.
 DR TIGRfam; TIGR00091; TIGR00091; 1.
 DR PROSITE; PS00092; N6_MTASE; UNKNOWN 1.
 KW Transferase; Methyltransferase; tRNA processing; Complete proteome.
 FT DOMAIN 1 306 HEMK.
 FT DOMAIN 307 524 tRNA (GUANINE-N(7)-)-METHYLTRANSFERASE.
 SQ SEQUENCE 524 AA; 60079 MW; 3B574D6CF7DFDC28 CRC64;
 Alignment Scores:

```

Pred. No.: 310 Length: 524
Score: 7.00 Matchee: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatch: 0
Query Match: 1.05% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x HEMK_RICCN (1-524)
QY 538 GAGCTGATACAGCAGATGCG 558
|||||
DB 473 G|U|E|U|L|E|G|I|N|G|I|N|A|N|G|I|Y 479

RESULT 46
CAP_YEAST
ID CAP_YEAST STANDARD; PRT; 526 AA.
AC P1755;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Adenyl1 cyclase-associated protein (CAP).
GN SRV2 OR CAP1 OR YNL138W OR N1210 OR N1838.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=90235281; PubMed=2184942;
RA Field J., Vojtek A., Ballaster R., Bolger G., Colicelli J.,
RA Ferguson K., Geert J., Kataoka T., Michaeli T., Powers S., Riggs M.,
RA Rodgers L., Wieland I., Wheland B., Wiggler M.;
RT "Cloning and characterization of CAP, the S. cerevisiae gene encoding
RT the 70 kd adenyl1 cyclase-associated protein.";
RL Cell 61:319-327(1990).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=90235282; PubMed=2158860;
RA Fedor-Chalken M., Deschenes R.J., Broach J.R.;
RT "SRV2, a gene required for RAS activation of adenylate cyclase in
RT yeast.";
RL Cell 61:329-340(1990).
[3]
RN RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=96109932; PubMed=8619318;
RA Mallet L., Bussereau F., Jacquet M.;
RT "A 43.5 kb segment of yeast chromosome XIV, which contains MFA2,
RT MEPR2, CAP/SRV2, NAM9, FBI1/PPR1/RBP1, MOM22 and CPT1, predicts an
RT adenosine deaminase gene and 14 new open reading frames.";
RL Yeast 11:1195-1209(1995).
-1- FUNCTION: The N-terminal domain binds to adenyl1 cyclase, thereby
enabling adenyl1 cyclase to be activated by upstream regulatory
signals, such as Ras. The C-terminal domain is required for normal
cellular morphology and growth control.
-1- SUBCELLULAR LOCATION: Cell membrane.
-1- SIMILARITY: Belongs to the CAP family.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its use
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC CC or send an email to license@isb-sdb.ch).
-----
DR EMBL, Z46843; CAA86887.1; -
DR EMBL, M58284; AAA63569.1; -
DR EMBL, M32663; AAA35094.1; -
DR EMBL, Z71414; CAA96020.1; -
DR PIR, A34896; A34896.
DR PIR, I142; I3-MAR-02.

```

```

DR   PDB: 1KO5; 16-JAN-02.
DR   Germonline; 143144; -.
DR   SGO; S0005082; SRVZ.
DR   GO; GO:0007265; P:RAS protein signal transduction; IMP.
DR   InterPro; IPR001837; CAP.
DR   InterPro; IPR006599; CAP.
DR   Pfam; PF01213; CAP; 1.
DR   SMART; SM00673; CAP; 2.
DR   PROSITE; PS01086; CAP_1; 1.
DR   PROSITE; PS01089; CAP_2; 1.
KM   Membrane; 3D-structure.
FT   DOMAIN   262      300
FT   DOMAIN   277      282
FT   DOMAIN   526      57521 MW; 0EB4D41205E2D464 CRC64;
SQ   SEQUENCE   526 AA; 57521 MW; 0EB4D41205E2D464 CRC64;

Alignment Scores:
Pred. No.:      310      Length:      526
Score:          7.00      Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:      1.09% Indels:      0
DB:               1      Gaps:      0

US-10-029-345A-108_COPY_532_2532 (1-2001) * CAP_YEAST (1-526)
QY      252 ACTGCATCAATGTCACACCTT 232
      |||||
DB      485 Thr1a1e1a5nvala5nleu 491

RESULT 47
PBB_SCHPO
ID   PBB_SCHPO          STANDARD;   PRT;   532 AA.
AC   060109; 09CA427;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alkaline phosphatase (EC 3.1.3.1).
GN   SPBC14F5.13C.
OS   Schizosaccharomyces pombe (fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomycetes.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kang S.-W., Lim C.-J.;
RT   "Characterization of alkaline phosphatase gene from
RT   Schizosaccharomyces pombe."
RL   Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21849401; Pubmed=11859360;
RX   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RX   Sgouros J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
RX   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RX   Collins M., Connor A., Cronin A., Davis P., Felwell T., Fraser A.,
RX   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RX   Holtova S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagele K.,
RX   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RX   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RX   Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RX   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RX   Skelton J., Simmonds N., Squares R., Squares S., Stevens K.,
RX   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RX   Woodward J., Voicakeert G., Aert R., Robben J., Grymoprez B.,
RX   Maltjens I., Vanstreets E., Rieger M., Schaefer M., Mueller-Auer S.,
RX   Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RX   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RX   Egger P., Zimmermann W., Medler H., Wambut R., Purrelle B.,
RX   Goffeau A., Cadieu E., Dreano S., Gloux S., Delaure V., Mottier S.,
RX   Galibert F., Aves S., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RX   Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,

```

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revelante J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
CC -1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC alcohol + phosphate.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (potential).
CC -1- SIMILARITY: Belongs to the alkaline phosphatase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: AF316541; AAK07179.1; -;
DR EMBL: AL023780; CAI9331.1; -;
DR PIR: T39459; T39459.
DR HSSP: P00634; 1A1A.
DR GeneDB_Spombe, SPB014F5.13c; -;
DR InterPro: IPR001952; Alk_phosphatase.
DR Pfam: PF00245; alk_phosphatase; 1.
DR PRINTS: PR00113; ALKPHPTASE.
DR ProDom: PD001868; Alk_phosphatase; 1.
DR SMART: SM00098; alkPc; 1.
DR PROSITE: PS00123; ALKALINE PHOSPHATASE; 1.
KW Hydrolyase; Zinc; Magnesium; Transmembrane; Signal-anchor.
FT TRANSMEM 31 48
FT ACT_SITE 114 114
FT CONFLICT 513 514 PS -> HC (IN REF. 1).
SQ SEQUENCE 532 AA; 58666 MW; 57A84A6926D545C CRC64;

Alignment Scores:
Pred. No.: 309 Length: 532
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.05% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x PPB_SCHPO (1-532)

QY 1547 ACACAGCTTCCTTTTCGACC 1567
ID TCE2_AVESEA STANDARD; PRT; 535 AA.
DB 168 ThrProLaserPheSerAla 174

RESULT 48
TCE1_AVESEA
ID TCE1_AVESEA STANDARD; PRT; 535 AA.
AC P40412;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE T-complex protein 1, epsilon subunit (TCP-1-epsilon) (CCT-epsilon)
DE (TCP-K19).
OS Avena sativa (Oat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Aveneae; Avena.
OX NCBI_TaxID=4498;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Dewi; TISSUE=coleoptile, and Mesocotyl;
RX MEDLINE=94085629; PubMed=7903257;
RA Ehmman B., Krenz M., Mummert E., Schaefer E.;
RT "Two tcp-1-related but highly divergent gene families exist in oat
RT encoding proteins of assumed chaperone function.";
RL FEBS Lett. 336:313-316(1993).

CC -1- FUNCTION: Molecular chaperone; assist the folding of proteins upon
CC ATP hydrolysis. Known to play a role, in vitro, in the folding of
CC actin and tubulin.
CC -1- SUBUNIT: Hetero-oligomeric complex of about 850 to 900 kDa that
CC forms two stacked rings, 12 to 16 nm in diameter.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the TCP-1-chaperonin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: X75777; CAAS3396.1; -;
DR PIR: S40461; S40461.
DR HSSP: P48425; 1A6D.
DR InterPro: IPR002194; Chaperonin TCP-1.
DR InterPro: IPR001844; Chaperonin Cpn60.
DR InterPro: IPR002423; Cpn60/TCP-1.
DR InterPro: IPR008950; GroEL-ATPase.
DR Pfam: PF00118; Cpn60_TCP1; 1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PROSITE: PS00750; TCP1_1; 1.
DR PROSITE: PS00751; TCP1_2; 1.
DR PROSITE: PS00995; TCP1_3; 1.
KW Chaperone; ATP-binding; Multigene family.
FT KMECH 535 AA; 58899 MW; 6A8F847CA891BC32 CRC64;
SQ SEQUENCE 535 AA; 58899 MW; 6A8F847CA891BC32 CRC64;

Alignment Scores:
Pred. No.: 309 Length: 535
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.05% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x TCE1_AVESEA (1-535)

QY 226 AACATGAGTTGACATTGAT 246
ID TCE2_AVESEA STANDARD; PRT; 535 AA.
DB 259 LysHisLysValAspIleAsp 265

RESULT 49
TCE2_AVESEA
ID TCE2_AVESEA STANDARD; PRT; 535 AA.
AC P54411;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE T-complex protein 1, epsilon subunit (TCP-1-epsilon) (CCT-epsilon)
DE (TCP-K36).
OS Avena sativa (Oat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Aveneae; Avena.
OX NCBI_TaxID=4498;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Dewi; TISSUE=coleoptile, and Mesocotyl;
RX MEDLINE=94085629; PubMed=7903257;
RA Ehmman B., Krenz M., Mummert E., Schaefer E.;
RT "Two tcp-1-related but highly divergent gene families exist in oat
RT encoding proteins of assumed chaperone function.";
RL FEBS Lett. 336:313-316(1993).
CC -1- FUNCTION: Molecular chaperone; assist the folding of proteins upon
CC ATP hydrolysis. Known to play a role, in vitro, in the folding of
CC actin and tubulin.
CC -1- SUBUNIT: Hetero-oligomeric complex of about 850 to 900 kDa that
CC forms two stacked rings, 12 to 16 nm in diameter.

```
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the TCP-1 chaperonin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75778; CAA5397.1; -.
CC PIR; S40462; S40462.
CC HSSP; P48424; IABD.
CC InterPro: IPR002194; Chaperonin TCP-1.
CC InterPro: IPR001844; Chaperonin Cpn60.
CC InterPro: IPR002423; Cpn60/TCP-1.
CC InterPro: IPR008950; GroEL-ATPase.
CC Pfam; PF00118; Cpn60 TCP1; 1.
CC PRINTS; PR00298; CHAPERONIN60.
CC PRINTS; PR00304; TCOMPLEXTCP1.
CC PROSITE; PS00750; TCP1_1; 1.
CC PROSITE; PS00751; TCP1_2; 1.
CC PROSITE; PS00955; TCP1_3; 1.
CC Chaperone; ATP-binding; Multigene family.
CC Chaperone; ATP-binding; Multigene family.
CC SEQUENCE 535 AA; 59034 MW; 02761226F5F17B81 CRC64;

Alignment Scores:
Pred. No.: 309 Length: 535
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.05% Indels: 0
DB: 1 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x TCE2_AVESa (1-535)

QY 226 AACATAGGTTGACATTGAT 246
Db 259 LysHsLysValAspLeaSp 265

RESULT 50
TCPE_ARATH STANDARD; PRT; 535 AA.
AC 004450;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE T-complex protein 1, epsilon subunit (TCP-1-epsilon) (CCT-epsilon).
GN ATIG24510 OR F21J9.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OC NCB1_TaxID=3702;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RC MEDLINE=21016719; PubMed=1130712;
RA Theologis A., Ecker J.R., Palm C.J., Bowmen C.L., Brooke S.Y.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooke S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Eegu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gull J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Mailli R., Marziani A.,
RA Millescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shin P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
```

```
RA Uterback T., Van Aken S., Vayenberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RA thaliana";
RL Nature 408:816-820(2000).
CC -1- FUNCTION: Molecular chaperone; assist the folding of proteins upon
CC ATP hydrolysis. Known to play a role, in vitro, in the folding of
CC actin and tubulin.
CC -1- SUBUNIT: Hetero-oligomeric complex of about 850 to 900 kDa that
CC forms two stacked rings, 12 to 16 nm in diameter.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: Belongs to the TCP-1 chaperonin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AC000103; AAF97977.1; -.
CC HSSP; P48424; IABD.
CC InterPro: IPR002194; Chaperonin TCP-1.
CC InterPro: IPR001844; Chaperonin Cpn60.
CC InterPro: IPR002423; Cpn60/TCP-1.
CC InterPro: IPR008950; GroEL-ATPase.
CC Pfam; PF00118; Cpn60 TCP1; 1.
CC PRINTS; PR00298; CHAPERONIN60.
CC PRINTS; PR00304; TCOMPLEXTCP1.
CC PROSITE; PS00750; TCP1_1; 1.
CC PROSITE; PS00751; TCP1_2; 1.
CC PROSITE; PS00955; TCP1_3; 1.
CC Chaperone; ATP-binding; Multigene family.
CC Chaperone; ATP-binding; Multigene family.
CC SEQUENCE 535 AA; 59384 MW; 1CB56343A3AF1DC3 CRC64;

Alignment Scores:
Pred. No.: 309 Length: 535
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.05% Indels: 0
DB: 1 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x TCPE_ARATH (1-535)

QY 226 AACATAGGTTGACATTGAT 246
Db 259 LysHsLysValAspLeaSp 265

Search completed: June 21, 2004, 13:10:22
Job time : 99 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: June 21, 2004, 12:59:47 ; Search time 102.5 Seconds

(without alignments)
12318.062 Million cell updates/sec

Title: US-10-029-345A-108_COPY_532_2532

Perfect score: 667
Sequence: 1 gttcgtatggccatgatgac.....tggaaatcattgagttctcc 2001

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Word size: 1

Total number of hits satisfying chosen parameters: 344724

Minimum DB seq length: 473

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODEL=frame+npz.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool_p/US10029345/runat.21062004.124615.5691/app_query.fasta.1.2183
-DB=SPTRMBL.25 -QFMT=fastan -SUFFIX=01igo.rspc -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=01igo -TRANS=human40.cdi
-LIST=75 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=473 -MAXLEN=200000000
-USER=US10029345_QCGN_1.1_112_0runat.21062004.124615.5691 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

SPTREMBL.25.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	472	70.8	665	4 Q96N49	Q96N49 homo sapien

2	400	60.0	662	4 Q81VT8	Q81VT8 homo sapien
3	371	55.6	616	4 Q8N5T1	Q8N5T1 homo sapien
4	67	10.0	660	11 Q920R2	Q920R2 mus musculus
5	64	9.6	677	11 Q99MG6	Q99MG6 mus musculus
6	48	7.2	622	11 Q99MG5	Q99MG5 mus musculus
7	20	3.0	625	4 Q86S88	Q86S88 homo sapien
8	20	3.0	665	11 Q7RS29	Q7RS29 mus musculus
9	13	1.9	476	5 Q9VH8	Q9VH8 mus musculus
10	13	1.9	476	5 Q46122	Q46122 drosophila
11	10	1.6	498	16 Q82Z00	Q82Z00 enterococcus
12	9	1.4	550	16 Q89P10	Q89P10 bradyrhizob
13	9	1.3	606	5 Q8ST18	Q8ST18 caenorhabdi
14	9	1.3	657	5 Q8ST19	Q8ST19 caenorhabdi
15	8	1.2	486	16 Q8BEN9	Q8BEN9 xanthomonas
16	8	1.2	495	4 Q9H6T2	Q9H6T2 homo sapien
17	8	1.2	495	4 Q96A99	Q96A99 homo sapien
18	8	1.2	497	16 Q8BT27	Q8BT27 lactobacilli
19	8	1.2	513	10 Q86S58	Q86S58 oryza sativ
20	8	1.2	513	10 Q7XEV9	Q7XEV9 oryza sativ
21	8	1.2	513	16 Q92M58	Q92M58 rhizobium m
22	8	1.2	515	10 Q9FX19	Q9FX19 arabidopsis
23	8	1.2	525	10 Q81CP6	Q81CP6 arabidopsis
24	8	1.2	525	10 Q9LR07	Q9LR07 arabidopsis
25	8	1.2	526	12 Q65756	Q65756 bluetongue
26	8	1.2	534	12 Q9Y0X9	Q9Y0X9 randid herpe
27	8	1.2	544	16 Q82GP4	Q82GP4 streptomyce
28	8	1.2	545	16 Q92ZH5	Q92ZH5 rhizobium m
29	8	1.2	555	12 Q55887	Q55887 human para
30	8	1.2	555	12 Q55888	Q55888 recombinant
31	8	1.2	559	16 Q8ABA4	Q8ABA4 bacteroides
32	8	1.2	569	16 Q7WG72	Q7WG72 bordetella
33	8	1.2	569	16 Q7W4P6	Q7W4P6 bordetella
34	8	1.2	573	16 Q7UQD7	Q7UQD7 rhodospirill
35	8	1.2	605	3 Q9P8M8	Q9P8M8 aspergillus
36	8	1.2	605	4 Q7Z6C5	Q7Z6C5 homo sapien
37	8	1.2	615	16 Q8D394	Q8D394 wigglewort
38	8	1.2	612	10 Q9ZUK4	Q9ZUK4 arabidopsis
39	8	1.2	677	4 Q8N4M1	Q8N4M1 homo sapien
40	8	1.2	682	10 Q85201	Q85201 arabidopsis
41	8	1.2	694	16 Q82K16	Q82K16 salmoneila
42	8	1.2	707	17 Q8U1R9	Q8U1R9 pyrococcus
43	8	1.2	711	12 P87583	P87583 chipmunk pa
44	8	1.2	727	16 Q8Y2P4	Q8Y2P4 anabaena sp
45	8	1.2	752	8 Q98505	Q98505 guillardi
46	8	1.2	757	2 Q9K1B0	Q9K1B0 porphyromon
47	8	1.2	759	12 Q91C22	Q91C22 ct virus. o
48	8	1.2	773	10 Q8L3X5	Q8L3X5 oryza sativ
49	8	1.2	789	16 Q8UGX6	Q8UGX6 agrobacteri
50	8	1.2	791	16 Q92RA3	Q92RA3 rhizobium m
51	8	1.2	849	16 Q8P8D5	Q8P8D5 xanthomonas
52	8	1.2	858	11 Q8K058	Q8K058 mus musculus
53	8	1.2	869	2 Q9AJ94	Q9AJ94 legionella
54	8	1.2	892	11 Q8OT06	Q8OT06 mus musculus
55	8	1.2	935	5 Q85MR0	Q85MR0 drosophila
56	8	1.2	1001	11 Q8BNG7	Q8BNG7 mus musculus
57	8	1.2	1037	16 Q7U036	Q7U036 rhodospirill
58	8	1.2	1042	16 Q82OR6	Q82OR6 nitrosomona
59	8	1.2	1065	5 Q815U3	Q815U3 plasmodium
60	8	1.2	1073	5 Q9W1T8	Q9W1T8 drosophila
61	8	1.2	1153	5 Q93598	Q93598 caenorhabdi
62	8	1.2	1194	10 Q9FKP1	Q9FKP1 arabidopsis
63	8	1.2	1222	16 Q87XX8	Q87XX8 pseudomonas
64	8	1.2	1308	5 Q01924	Q01924 caenorhabdi
65	8	1.2	1717	11 Q7RS29	Q7RS29 mus musculus
66	8	1.2	1762	10 Q94DC2	Q94DC2 oryza sativ
67	8	1.2	1885	17 Q8TOD1	Q8TOD1 mechanosarc
68	8	1.2	1900	12 Q9W3P1	Q9W3P1 sindbis-lik
69	8	1.2	1901	12 Q88429	Q88429 sindbis-lik
70	8	1.2	2487	12 Q9YJX8	Q9YJX8 sindbis vir
71	8	1.2	2500	12 Q88431	Q88431 sindbis-lik
72	8	1.2	2513	12 Q87644	Q87644 sindbis vir
73	8	1.2	2517	12 Q9YJX6	Q9YJX6 sindbis-lik
74	8	1.2	2761	5 Q16447	Q16447 caenorhabdi

C 75 8 1.2 3972 16 Q9S0R8 Q9S0R8 streptomyce

ALIGNMENTS

RESULT 1

Q96N49 PRELIMINARY; PRT; 665 AA.
ID Q96N49
AC Q96N49;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ1411.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SOURCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsu T., Sato H., Wakamatsu A., Ishii S., Yamamoto Y., Isono Y.,
RA Kawai-Hio Y., Saigo K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Magatsuna M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isegaki T.;
RT "MEDO human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK055973; BAB71060.1; -.
DR HSSP; Q16828; IMKP.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR001763; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR PRINTS; PRO1764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS0206; RHODANASE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
DR Hypothetical protein; Hydrolase.
KW SEQUENCE 665 AA; 73058 MW; 1EABDDF08460DF CRC64;
SQ

Alignment Scores:

Pred. No.: 0 Length: 665
Score: 472.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.76% Indels: 0
Dbs: 4 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x Q96N49 (1-665)

QY 586 ACCTGTCGAAGCTGACTTATCCCGAGTCTCATTTCTGCGCTGCTGGAATGAC 645
DB 194 ThrcyprolysprroahpnhellProgluSerhisshelenuAargValProvalAsnasp 213
QY 646 AGCTTTGTGAGAAAATTTTCCGCTGTTGGACAATCAGTAGATTTCATTGAGAAAGCA 705
DB 214 SerPheCysGluysIleLeuProTirPheusAplysSerValaAapPheilleGluysala 233
QY 706 AAAGCCTCCATGATGATGTTCTAGTGACATGTTAGCTGGAGATCTCCGCTCCGACAC 765
DB 234 LysAlaSerAsnGlyCyValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThr 253
QY 766 ATCGCTATGCGCTACATCATGAAAGATGAGCATGTCTTATGATGAAGCTTACAGATT 825

DB 254 lLeallleallatYrillewetylsArgwetasPheSerleuAsnspGluAlaYrArgPhe 273
QY 826 GTGAAAAGAAAAAGACCTACTATATCTCCAACTCAATTTTCTGGGCCAATCTCTGGAC 885
DB 274 VallysgIuysArgProthrIleSerProAsnPhenAsnPhenleuGlyInleuLeuAsp 293
QY 886 TATGAGAAAGAAATTAAAGAACCAAGTGGAGCATCGGGCCAAAGCAAACTCAAGCTG 945
DB 294 TyrgIuysIysIleuysAsnGlnthrgYlaSeryIleProlySeryIleuysleu 313
QY 946 CTGACCTGGAGAAAGCAAAATGAACCTGTCTCTGTCTTCAGAGGGTGGACAAAGC 1005
DB 314 LeuHisleuGluysProAsnGluProvalProAlaValSerGluGlyGlnlySer 333
QY 1006 GAGAGCCCTCACTGCTCACTCTGTCGCACTGTCTGCTTCAAGAGGACAGAGCAAAAG 1065
DB 334 GluthrProleuSerProProCyAlaAspSerValThrSerGluAlaAlaGlyGlnArg 353
QY 1066 CCGCTGCATCCCGGACAGCGTCCAGGCTGCCAGCGTGCAGCGCTGTTAGAGAC 1125
DB 354 ProvalHisProAlaSerValProSerValProSerValGlnProSerleuLeuAsp 373
QY 1126 AGCCCGCTGTACAGAGCGCTCAGTGGCTGCACCTGTCCGACAGAGCTGGAAAGACAC 1185
DB 374 SerProleuValGlnAlaLeuSeryGlyLeuHisIleuSerAlaAspArgLeuGluAspSer 393
QY 1186 AATTAAGCTCAAGCGTCTCTCTCTCTGATATCAAAACAGTTTCATATTCAGCCAGCATG 1245
DB 394 AsnlySleuysAsArgSerPheSerleuAspIlelySeryValSeryIleSerAlaSerMet 413
QY 1246 GCAGATCTTACATGAGGCTTCTCTCTCTCAGAAAGATCTTGAATATCAAACTTCC 1305
DB 414 AlaAlaSerleuHisGlyIlePheSerSerSerGluAsnAlaLeuGluYrYrlyProSer 433
QY 1306 ACTACTGTGATGGAGACCAACAGCTATGCCAGTTTCTCCCTGTTCCAGAACTTCCGAG 1365
DB 434 ThrThrleuAsnArglyThrAsnlySleuysGlnPheSerProvalGlnGluLeuSeryGlu 453
QY 1366 CAGATCCCGAAACACAGCTCTGATAGAGAGAGAGACATCCCAAGAACTCCAGAC 1425
DB 454 GlnThrProGluThrSerProAspSeryGluGluAlaSerIleProlySeryleuGlnThr 473
QY 1426 GCCAGCGCTTACAGACGACGACGAGCAAGCATTTGATGATGATGATGATGATGATGATG 1485
DB 474 AlaArgProSerAspSerGlnSeryIleSeryIleSeryIleSeryIleSeryIleSeryIle 493
QY 1486 ACCGCTCAGAGGCTCCCTTTTATCTCCACTGCATGAACTGGAGAGTGGAGAGCAATTAC 1545
DB 494 ThrAlaGlnArgSeryleuLeuSeryProleuHisIleArgSeryGlySeryValGluAsnIle 513
QY 1546 CACACAGCTTCTTTTGGCGCTTCCACAGCCAGCAGCAGCACTTCAAGCAAGCTTGGCTGC 1605
DB 514 HisThrSerPheleuPheGlyleuSerThrSeryGlnGlnHisIleuThrlySeryAlaGly 533
QY 1606 CTGGGCTTAAAGGGCTGGCACTCGGATATCTTGGCTCCCGCCAGACCTTACCTTCCCTG 1665
DB 534 LeuGlyleuysGlyTirPheIleSerAspIleleuAlaProGlnThrSerThrProSerleu 553
QY 1666 ACCAGCAGCTGGTATTTTGGCCACAGAGTCTTCAAGCTTCACTGCTGCTCAGCCATCAC 1725
DB 554 ThrSerSerTirPlyPheAlaThrGluSerSerHisSeryIleSeryIleSeryIleSeryIle 573
QY 1726 GGAGGAGTGCAGTAACTTCTGCTTCAAGCTTGCAGCCAGCTGCGCATTTGCGAGACCA 1785
DB 574 GlyIlyserAlaSerTlySeryIleSeryIleSeryIleSeryIleSeryIleSeryIle 593
QY 1786 GTCTATTCTGCGGAGGCGGAGCAAGCAAGTGAACAAGCTGATTCGCGGAGGCTGG 1845
DB 594 ValIlyserValArgArgArgGlnlySproSerAspArgAlaAspSerArgArgSerTirp 613
QY 1846 CATGAAGAGAGCCCTTGAAGCAAGTATTAACCAAGAGCTGCAAAATTTTGA 1905

ID Q920R2 PRELIMINARY; PRT; 660 AA.
AC Q920R2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MAP kinase phosphatase-7.
OS Mus musculus (Mouse).
GN DUSP16 OR 3830417M17RIK OR MKP-7.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=21486429; PubMed=11489891;
RA Maeda K., Shima H., Watanabe M., Kikuchi K.,
RT "MKP-7, a Novel Mitogen-Activated Protein Kinase Phosphatase,
Functions as a Shuttle Protein".
RL J. Biol. Chem. 276:39002-39011(2001).
DR EMBL: AB052157; BAB47240.1; -.
DR HSSP: Q16828; 1MKP.
DR MGD: MG1:1917936; C:cytoplasm; IDA.
DR GO: GO:0005737; C:cytoplasm; IDA.
DR GO: GO:0005634; C:nucleus; IDA.
DR GO: GO:0005515; F:protein binding; IPI.
DR GO: GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS phosphatase.
DR InterPro: IPR008343; MAPK phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANSE_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydroxylase; Kinase.
SQ SEQUENCE 660 AA; 72695 MW; DB609FCDAD4AA309 CRC64;
Alignment Scores:
Pred. No.: 1,14e-60 Length: 660
Score: 67.00 Matches: 67
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.04% Indels: 0
Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x Q920R2 (1-660)
QY 49 TTGGTGGCTCTGCTGGAAGTGAACGAAAGTCTGTAATGATACCGCCATT 108
DB 15 LeuValAlaIleuIeuGluSerGlyThrGluValIleuIeuIleAspSerArgProPhe 34
QY 109 GTGGAATACATCATCTCCACATTTTGAAGCCATTAAATCACTGCTCCAACTTATG 168
DB 35 ValGluTyrAsnThrSerHisIleLeuGluAlaIleAsnIleAsnGlySerIleuMet 54
QY 169 AACGGAAGTTGACGACGACGAAGTGTAAATTAACAGAGCTCATCCAGCACTTCAGGAAA 228
DB 55 LysArgArgLeuGlnGlnIleuIleuValIleuIleuIleuIleuIleuIleuIleuIleuIleu 74
QY 229 CATAGGTTGACATGATGTC 249
DB 75 HisIleValaIleAspIleAspCys 81
RESULT 5
ID Q99MG6 PRELIMINARY; PRT; 677 AA.
AC Q99MG6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Map kinase phosphatase-M A1 isoform.
GN DUSP16 OR 3830417M17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Matsuguchi T., Musikacharen T., Johnson T.R., Kraft A.S.,
RT Yoshikai Y.;
RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
Activation in Macrophages".
RL Mol. Cell. Biol. 20:6999-7009(2001).
DR EMBL: AF345951; AAK35052.1; -.
DR HSSP: Q16828; 1MKP.
DR MGD: MG1:1917936; C:cytoplasm; IDA.
DR GO: GO:0005737; C:cytoplasm; IDA.
DR GO: GO:0005634; C:nucleus; IDA.
DR GO: GO:0005515; F:protein binding; IPI.
DR GO: GO:0000188; P:inactivation of MAPK; IDA.
DR InterPro: IPR000340; DS phosphatase.
DR InterPro: IPR008343; MAPK phosph.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS50206; RHODANSE_3; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydroxylase; Kinase.
SQ SEQUENCE 677 AA; 74550 MW; 8B6D5B7096CEC2FC CRC64;
Alignment Scores:
Pred. No.: 1,63e-57 Length: 677
Score: 64.00 Matches: 64
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.60% Indels: 0
Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x Q99MG6 (1-677)
QY 346 AGCTTAACCTCTGTTCACTGCTTGAGGTGGTTCGATGCTCTGCTTCTTCCT 405
DB 114 SerPheAsnSerValHisIleuIleuAlaGlyGlyPheAlaGluPheSerArgCysPhePro 133
QY 406 GGCTCTGTAAGAAATCACTAGTCCCTACCTGATTTCTGAGCTTGCCTTAACCT 465
DB 134 GlyLeuCysGlnGlyLysSerThrIleuValProThrCysIleSerGlnProCysLeuPro 153
QY 466 GTTGCACATTTGGGCAACCCGAATCTTCCCAATCTTTATCTTGGCTGCAGCAGAT 525
DB 154 ValAlaAsnIleGlyProThrArgIleuProAsnLeuTyrLeuGlyCysGlnArgAsp 173
QY 526 GTTCCTCAACAG 537
DB 174 ValIleuAsnLys 177
RESULT 6
ID Q99MG5 PRELIMINARY; PRT; 622 AA.
AC Q99MG5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Map kinase phosphatase-M A2 isoform.

GN DUSP16 OR 3830417M17RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
 RA Yoshikai Y.
 RA "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
 RT Activation in Macrophages."
 RT Mol. Cell. Biol. 20:6999-7009(2001).
 DR EMBL: AF345952; AAK35053.1; -
 DR HSSP: O16828; 1MKP.
 DR MGD: MGI:1917936; DUSP16.
 DR GO: GO:0005737; C:cytoplasm; IDA.
 DR GO: GO:0005634; C:nucleus; IDA.
 DR GO: GO:0005515; F:protein binding; IPT.
 DR GO: GO:0000188; P:inactivation of MAPK; IDA.
 DR InterPro: IPR000340; DS phosphatase.
 DR InterPro: IPR008343; MAPK phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR PRINTS: PRO1764; MAPKPHPTASE.
 DR SMART: SM00450; RHOD; 1.
 DR SMART: SM00195; DSPC; 1.
 DR PROSITE: PS50206; RHODANSE_3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolase; kinase.
 SQ SEQUENCE 622 AA; 68672 MW; 6C0CA4EEB909B98 CRC64;

Alignment Scores:
 Pred. No.: 1.12e-40 Length: 622
 Score: 48.00 Matches: 48
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.20% Indels: 0
 DB: 11 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q99WGS (1-622)

OY 586 ACCTGTCGAAGCTGACTTATCCCGAGTCTCATTTCTGCGTGTGCTGATGAC 645
 |||||
 DB 139 Thrcysprolyspromaspheilleprogluserhisphleuargvalprovalasnap 158
 |||||

OY 646 AGCTTTTGTGAGAAATTTTCCGCTGTGACAAATCAGTATTTTCATTGAGAAAGCA 705
 |||||
 DB 159 Serphecysgilylileleuprotipheuasphysservalasphelleglulysala 178
 |||||

OY 706 AAAGCTCCATGATGATGTTCTA 729
 |||||
 DB 179 Lysalaseramnglycyvalleu 186
 |||||

RESULT 7
 086SS8 PRELIMINARY: PRT; 625 AA.
 AC 086SS8:
 DT 01-UN-2003 (TREMBLrel. 24, Created)
 DT 01-UN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Similar to dual specificity phosphatase 8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, and Astrocytoma;

RA Strausberg R.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC045110; AAF45110.1; -
 DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro: IPR000340; DS phosphatase.
 DR InterPro: IPR008343; MAPK phosph.
 DR InterPro: IPR002965; P:rich exten.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR PRINTS: PRO1764; MAPKPHPTASE.
 DR PRINTS: PRO1217; PRICHEXTENSN.
 DR SMART: SM00195; DSPC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS50206; RHODANSE_3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 SQ SEQUENCE 625 AA; 65826 MW; C7C808407B724FPC CRC64;

Alignment Scores:
 Pred. No.: 3.24e-11 Length: 625
 Score: 20.00 Matches: 20
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.00% Indels: 0
 DB: 4 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q86SS8 (1-625)

OY 730 GTGCACGTGTTAGTGGATCTCCGCTCCGACCAATGCTATGCTTACATCATGAG 789
 |||||
 DB 244 Valhscysleualaglyliserargserlathrlleallalealalylllemelys 263
 |||||

RESULT 8
 07TS29 PRELIMINARY: PRT; 665 AA.
 ID 07TS29:
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Dusp8 protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Martusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,
 RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carinici P., Bouffard G.G.,
 RA Raba S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Wai M.A.,
 RT Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Straubeberg R.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC052705; AAF52705.1; -
 SQ SEQUENCE 665 AA; 69021 MW; 9166E36A8835249F CRC64;

Alignment Scores:
 Pred. No.: 3 21e-11 Length: 665
 Score: 20.00 Matches: 20
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.00% Indels: 0
 DB: 11 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x Q7TSZ9 (1-665)

QY 730 GTCACCTGTTTACCTGGATCTCCCGCTCCGCCACCATGCTATCGCTACATCATGAAG 789
 |||||
 DB 244 VALHICYLEULAGIYIESERARGSERALATHRIEALIEALIEATYRIEMETLYS 263

RESULT 9
 QYHV8 PRELIMINARY; PRT; 476 AA.

AC Q9VHV8;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE PUC protein (SD08157P).
 PUC OR CG7850.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX MEDLINE=7227;
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman J.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brodeur P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferreira C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Giordano A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jaitai M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Juelich B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mikhlin N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puft V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slater E., Spalding A.C., Stapleton M., Strong R., Sun E.,
 RA Strebner R., Tector C., Turner R., Venter J., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasekman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang R., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003677; AAF54191.1; -
 DR EMBL: AY061616; AAL29164.1; -
 DR HSSP: Q16828; 1MKP.
 DR FlyBase: FBgn0004210; puc.
 DR GO: GO:0008759; F:JUN kinase phosphatase activity; IMP.
 DR GO: GO:0017017; F:MAP kinase phosphatase activity; NAS.
 DR GO: GO:0030036; P:actin cytoskeleton organization and biogenesis; IMP.
 DR GO: GO:0046843; P:dorsal appendage formation; IMP.
 DR GO: GO:0008544; P:epidermal differentiation; IMP.
 DR GO: GO:0007254; P:JNK cascade; IMP.
 DR GO: GO:0046844; P:microtubule formation; IMP.
 DR GO: GO:0046329; P:negative regulation of JNK cascade; NAS.
 DR GO: GO:0016318; P:maternal rotation; NAS.
 DR GO: GO:0006470; P:protein amino acid dephosphorylation; IDA.
 DR GO: GO:0007396; P:wound healing; NAS.
 DR GO: GO:0042060; P:wound healing; IFP.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR SMART: SM00195; DSPC; 1.
 DR PROSITE: PS00183; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS00056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS0054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolyase.
 SQ SEQUENCE 476 AA; 51270 MW; 8E5D281349F0F4F CRC64;

Alignment Scores:
 Pred. No.: 0.000778 Length: 476
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.95% Indels: 0
 DB: 5 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x Q9VHV8 (1-476)

QY 742 GCTGGATCTCCGCTCCGCCACCATGCTATCGCTTAC 780
 |||||
 DB 218 ALAGIYIIESERARGSERALATHRIEALIEALIEATYRI 230

RESULT 10
 ID 046122
 AC 046122; PRELIMINARY; PRT; 476 AA.
 DT 01-JUN-1998 (TEMBLrel. 06, Created)
 DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE PUCKERED protein.
 GN PUC OR PUCKERED OR CG7850.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX MEDLINE=7227;
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman J.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brodeur P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferreira C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Giordano A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jaitai M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Juelich B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mikhlin N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puft V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slater E., Spalding A.C., Stapleton M., Strong R., Sun E.,
 RA Strebner R., Tector C., Turner R., Venter J., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasekman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang R., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RA Chisee S.;
RT "The sequence of C. elegans cosmid F08B1."
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U23178; AAK68300.1; -.
DR HSSP; Q16828; IMKP.
DR WormRep; F08B1.1b; CE27918.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008138; F:protein tyrosine/serine/threonine phosphata. . .; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hypothetical protein; Hydrolase.
SQ SEQUENCE 606 AA; 64962 MW; 5B71795C6ED58058 CRC64;
KW
Alignment Scores:
Pred. No.: 12 2 Length: 606
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 5 Gaps: 0
US-10-029-345a-108_COPY_532_2532 (1-2001) x Q8ST18 (1-606)
QY 733 CACTGTTAGCTGGATCTCCGCTCC 759
Db |||||
210 HiscYleuAlaGlyIleSerArgSer 218
RESULT 14
Q8ST19 PRELIMINARY; PRT; 657 AA.
AC Q8ST19;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F08B1.1a.
GN F08B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN=Briscot N2;
RA Chisee S.;
RT "The sequence of C. elegans cosmid F08B1."
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U23178; AAK68299.1; -.
DR HSSP; Q16828; IMKP.
DR WormRep; F08B1.1a; CE27918.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hypothetical protein; Hydrolase.
SQ SEQUENCE 657 AA; 71002 MW; A0D9153DE6326B43 CRC64;
KW
Alignment Scores:
Pred. No.: 12 Length: 657
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 5 Gaps: 0
US-10-029-345a-108_COPY_532_2532 (1-2001) x Q8ST19 (1-657)
QY 733 CACTGTTAGCTGGATCTCCGCTCC 759
Db |||||
261 HiscYleuAlaGlyIleSerArgSer 269
RESULT 15
Q8PEN9
ID Q8PEN9 PRELIMINARY; PRT; 486 AA.
AC Q8PEN9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Outer membrane efflux protein.
GN NODT OR XAC4306.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
Camarote G., Canavan F., Cardozo J., Chamberg F., Ciapina L.P.,
Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
Racaiyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
Martins E.C., Melandri J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,

RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezsa R.I.D.,
RA Spradode dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Serubal J.C., Kitajima J.P.,
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463 (2002).
DR EMBL; AB012084; AAM9136.1; -
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003423; OEP.
DR InterPro; IPR004377; Prok_lipoprot_S.
DR Pfam; PF02321; OEP; 2.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN, 1.
KW Complete proteome.
SQ SEQUENCE 486 AA; 51711 MW; 7E0D7E71A18F02D CRC64;

Alignment Scores:
Pred. No.: 141 Length: 486
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 16 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x Q8P6N9 (1-486)

QY 103 GCCGGTCATCAATTAGCAGCACTT 80
DB 253 AAGGlyTyrGlnMetuAlaAlaLeu 260

RESULT 16

Q9H6T2 PRELIMINARY; PRT; 495 AA.
ID Q9H6T2
AC Q9H6T2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ21915.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEO human cDNA sequencing project";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK025568; BAB15171.1; -
DR GO; GO:0003899; P:DNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
KW Hypothetical protein.
SQ SEQUENCE 495 AA; 56749 MW; F30E6916AD9BED3 CRC64;

Alignment Scores:
Pred. No.: 141 Length: 495
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 4 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x Q9H6T2 (1-495)

QY 386 AGTTCTCTGTTGTTTCCCTGACC 409
DB 316 SerSerLeuValValSerLeuAla 323

RESULT 17

Q96AG9 PRELIMINARY; PRT; 495 AA.
ID Q96AG9
AC Q96AG9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue-Skin;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017115; AAH17115.1; -
DR GO; GO:0003899; P:DNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
KW Hypothetical protein.
SQ SEQUENCE 495 AA; 56750 MW; 2D7AE1D91E5D7871 CRC64;

Alignment Scores:
Pred. No.: 141 Length: 495
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 4 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x Q96AG9 (1-495)

QY 386 AGTTCTCTGTTGTTTCCCTGACC 409
DB 316 SerSerLeuValValSerLeuAla 323

RESULT 18

Q88T27 PRELIMINARY; PRT; 497 AA.
ID Q88T27
AC Q88T27;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Threonine synthase (EC 4.2.99.2).
GN THRC OR LP 2758.
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OC NCBI_TaxID=1590;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NCIMB 8826 / WCPS1;
RX MEDLINE=22480296; PubMed=12566566;
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hofer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ureing B.,
RA De Vos W.M., Stenzen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCPS1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995 (2003).
DR EMBL; AL935260; CAD64989.1; -
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR001926; B6_enzyme_beta.
DR InterPro; IPR000634; S/T_dehydrtse_BS.
DR Pfam; PF00291; PALP; 1.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW Lyase; Complete proteome.
SQ SEQUENCE 497 AA; 54569 MW; DA4344785D0275CC CRC64;

Alignment Scores:
Pred. No.: 141 Length: 497
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 4 Gaps: 0

Pred. No.: 141 Length: 497
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 16 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x Q86R27 (1-497)

QY 1733 GTGCCACTTACTCTGCTACAGCT 1756

DB 34 ValProValThrLeuProThrAla 41

RESULT 19

Q86S68 PRELIMINARY; PRT; 513 AA.

AC Q86S68; 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DE Putative transposable element.

GN OSJNB0004A10.10.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=4530;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=Nipponbare;

RA McCombie W.R., de la Bastide M., Spiegel L., Nascimento L., Ballja V.,

RA Zlatavorn T., Bell M., Preston R., Kirchoff K., Kuit K., Baker J.,

RA Santos L., Miller B., Cumins D.M., Katzenberger F., Muller S.,

RA Shah R., King L., Yang C., Dike S., O'Shaughnessy A., Palmer L.,

RA Dechha N.;

RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone

RT OSJNB0004A10, from chromosome 10, complete sequence.";

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC098682; AAM01106.1; -.

DR Gramene; Q86S68; -.

DR InterPro; IPR008906; HATC.

DR Pfam; PF05699; hATC; 1.

SQ SEQUENCE 513 AA; 58087 MW; 978805247396E6AA CRC64;

Alignment Scores:

Pred. No.: 140 Length: 513

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.24% Indels: 0

DB: 10 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x Q86S68 (1-513)

QY 637 CAGGACACGACGAGAAATGAGACT 614

DB 490 GlnAlaHisAlaGlyAsnGlnThr 497

RESULT 20

Q7XFX9 PRELIMINARY; PRT; 513 AA.

AC Q7XFX9; 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DE Putative transposable element.

GN OSJNB0004A10.10.

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RP STRAIN=cv. Nipponbare;

RA The Rice Chromosome 10 Sequencing Consortium;

RT "In-depth view of structure, activity, and evolution of rice

RT Chromosome 10.";

RT Science 300:1566-1569(2003).

RN [2] SEQUENCE FROM N.A.

RP STRAIN=cv. Nipponbare;

RA Buell C.R., Wang R.A., McCombie W.R., Messing J., Yuan Q.;

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB017075; AAP52942.1; -.

SQ SEQUENCE 513 AA; 58087 MW; 978805247396E6AA CRC64;

Alignment Scores:

Pred. No.: 140 Length: 513

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.24% Indels: 0

DB: 10 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x Q7XFX9 (1-513)

QY 637 CAGGACACGACGAGAAATGAGACT 614

DB 490 GlnAlaHisAlaGlyAsnGlnThr 497

RESULT 21

Q92M58 PRELIMINARY; PRT; 513 AA.

AC Q92M58; 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DE Putative sigma-54-dependent transcription regulator protein.

GN TACA OR R02795 OR SMC04011.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=1021.

RA MEDLINE=21396507; PubMed=11481430;

RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,

RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,

RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,

RA Pohl T., Portetiere D., Puhler A., Purnelle B., Rampeger U.,

RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;

RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

DR EMBL; AL591791; CAC47374.1; -.

DR HSSP; P06143; 1D4Z.

DR GO; GO:0000166; F:nuclotide binding; IEA.

DR GO; GO:0003700; F:transcription factor activity; IEA.

DR GO; GO:0000156; F:two-component response regulator activity; IEA.

DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR002197; HTH_Fis.

DR InterPro; IPR001789; Response_reg.

DR InterPro; IPR008298; Reg_reg_Ntc.

DR Pfam; PF02954; HTH_8; 1.

DR Pfam; PF00072; response_reg; 1.

DR Pfam; PF00158; Sigma54_activat; 1.

DR PRINTS; PR01590; HTHRFS.

DR ProDom; PD000039; Response_reg; 1.

DR SMART; SM00382; AAA; 1.

DR SMART; SM00448; REC; 1.

DR TIGRFAMs; TIGR01199; HTH_fis; 1.

DR PROSITE; PSS0110; RESPONSE_REGULATORY; 1.

DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
 DR PROSITE; PS00676; SIGMA54_INTERACT_2; 1.
 DR PROSITE; PS00688; SIGMA54_INTERACT_3; 1.
 DR PROSITE; PS50045; SIGMA54_INTERACT_4; 1.
 DR PIRSF; PIRSF003187; Res_Reg_Ntrc; 1.
 KW Complete proteome.
 SQ SEQUENCE 513 AA; 55902 MW; 5E1964533DE8F65 CRC64;

Alignment Scores:

Pred. No.:	140	Length:	513
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.24%	Indels:	0
DB:	16	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q92ME8 (1-513)

QY 1776 GCAAGTGGGACGCTGCTGACGCT 1753

Db 417 AlaSerGlyGlnLeuAlaAlaAla 424

RESULT 22

Q9FX19 PRELIMINARY; PRT; 515 AA.

AC Q9FX19; 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE F6P9.1 Protein (Altig19940/F6P9_1).
 GN F6P9.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;

SEQUENCE FROM N.A.

RA Federpiele N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
 RA Alafafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
 RA Buchler E., Chao Q., Chin C., Chou J., Choi E., Gonzalez A.,
 RA Hwang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
 RA Lenz C., Liu A., Liu S., Mukharasy N., Pham P., Sakano H., Shinn P.,
 RA Toriumi M., Vayberg M., Yu G., Ecker J., Theologis A., Davis R.W.,
 RL Submitted (SSP-2000) to the EMBL/GenBank/DBJ databases.
 [2]

SEQUENCE FROM N.A.

RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Gibson H.A.,
 RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J.,
 RA Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
 RA Davis R.W., Theologis A., Ecker J.R.,
 RT "Arabidopsis cDNA clones."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 [3]

SEQUENCE FROM N.A.

RA Kim C.J., Chen H., Cheuk R., Shinn P., Banh J., Bowser L.,
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
 RT "Arabidopsis ORF clones."

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC007797; AAG12562.1; -

DR EMBL; AY048245; AAK82507.1; -

DR PIR; G66332; G66332. -

DR HSP; P26221; ITF4.

DR GO; GO:0004553; F:hydrolyase activity, hydrolyzing O-glycosyl . . .; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR001701; Glyco_hydro_9.
 DR InterPro; IPR008928; Glyco_trans_6np.
 DR Pfam; PF00759; Glyco_hydro_9; 1.
 DR PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.
 SQ SEQUENCE 515 AA; 56707 MW; F21BEC7479B5E209 CRC64;

Alignment Scores:

Pred. No.:	140	Length:	515
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	10	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q9FX19 (1-515)

QY 374 GTGGGTTTCTGAGTTCTTCGTT 397

Db 499 ValGlyLeuLeuSerSerLeuVal 506

RESULT 23

Q8LCP6 PRELIMINARY; PRT; 525 AA.

AC Q8LCP6; 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Endo-beta-1,4-glucanase, putative.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;

SEQUENCE FROM N.A.

RA Haas B.J., Volkovsky N., Town C.D., Troughan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.,
 RT "Full-length messenger RNA sequences greatly improve genome
 annotation."
 RL Genome Biol. 0:0-0(2002).
 [2]

SEQUENCE FROM N.A.

RA Brover V., Troughan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.,
 RT "Full-length cDNA from Arabidopsis thaliana."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY086475; AAM63477.1; -
 DR GO; GO:0004553; F:hydrolyase activity, hydrolyzing O-glycosyl . . .; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR001701; Glyco_hydro_9.
 DR InterPro; IPR008928; Glyco_trans_6np.
 DR Pfam; PF00759; Glyco_hydro_9; 1.
 DR PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.
 SQ SEQUENCE 525 AA; 57915 MW; 26D7241334F5E32A CRC64;

SEQUENCE FROM N.A.

RA Kim C.J., Chen H., Cheuk R., Shinn P., Banh J., Bowser L.,
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
 RA Davis R.W., Theologis A., Ecker J.R.,
 RT "Arabidopsis ORF clones."

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC007797; AAG12562.1; -

DR EMBL; AY048245; AAK82507.1; -

DR PIR; G66332; G66332. -

DR HSP; P26221; ITF4.

ID	Accession	Source	Length	Score
AD	09LR07	PRELIMINARY, PRF, 525 AA.	140	8.00
DT	01-OCT-2000	(TREMBLrel. 15. Created)		
DT	01-OCT-2000	(TREMBLrel. 15, last sequence update)		
DT	01-OCT-2003	(TREMBLrel. 25, last annotation update)		
DE	PI045.13	(Hypothetical protein) (Putative		
DE	endo-beta-1,4-glucanase).			
CN	PI045.13	OR AT1G75680.		
CS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.			
OX	NCBI_TaxID=3702;			
RN				
RP	SEQUENCE FROM N.A.			
RA	Chen R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,			
RA	Khan S., Kim C., Altafi H., Bai B., Chin C., Chlou J., Choi E.,			
RA	Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,			
RA	Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharbeky N.,			
RA	Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,			
RA	Thavert A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,			
RA	Theologis A., Ecker J.;			
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.			
RN				
RP	SEQUENCE FROM N.A.			
RA	Chen R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,			
RA	Khan S., Kim C., Altafi H., Bai B., Chin C., Chlou J., Choi E.,			
RA	Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,			
RA	Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharbeky N.,			
RA	Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,			
RA	Thavert A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,			
RA	Theologis A., Ecker J.;			
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
RN				
RP	SEQUENCE FROM N.A.			
RA	Yanada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,			
RA	Dale J.M., Gibson H.A., Goldenrich A.D., Jjing P.X., Lee J.M.,			
RA	Quach H.L., Tang C., Toriumi M., Yu G., Bowser L., Carninci P.,			
RA	Chen H., Chen R., Hayashizaki Y., Ishida J., Jones T., Kamita A.,			
RA	Karlin-Neumann G., Kawai J., Kim C., Kosemura E., Lam B., Lin J.,			
RA	Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,			
RA	Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,			
RA	Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;			
RT	"Full Length cDNA of gene F1045.13 (GI:9365363)." ;			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
RN				
RP	SEQUENCE FROM N.A.			
RA	Yanada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,			
RA	Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,			
RA	Yu G., Yuan S., Carninci P., Chen H., Chen R., Hayashizaki Y.,			
RA	Ishida J., Jones T., Kamita A., Kawai J., Kim C.J., Narusaka M.,			
RA	Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,			
RA	Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,			
RA	Theologis A.;			
RT	"Arabidopsis Open Reading Frame (ORF) Clones." ;			
RL	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AC006434; AAF87112.1; -			
DR	EMBL; AY039938; AAK64042.1; -			
DR	EMBL; AY150451; AAN12892.1; -			
DR	PIR; E96786; E96786.			
DR	HSSP; P26221; ITH4.			
DR	GO; GO:0004553; F:hydrolyase activity, hydrolyzing O-glycosyl . . ; IEA			
DR	GO; GO:0005975; P:carbohydrate metabolism; IEA.			
DR	InterPro; IPR001701; Glyco_hydro_9.			
DR	InterPro; IPR008928; Glyco_trans_6np.			
DR	Pfam; PF00759; Glyco_hydro_9; 1.			
DR	PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.			
KW	Hypothetical protein.			
SQ	SEQUENCE 525 AA; 57868 MW; FBC3704F4B8B871 CRC64;			

```

US-10-029-145A-108_COPY_532_2532 (1-2001) x Q9LR07 (1-525)
QY 374 GTGGGTTTGTGACTGTTCTGCTT 397
DB 508 ValGlyLeuSerLeuSerLeuVal 515

RESULT 25
065756 PRELIMINARY; PRT; 526 AA.
ID Q65756
AC Q65756;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VP5.
OS Bluetongue virus.
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OC NCBI_TaxID=40051;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype 11 strain UC-8;
RA Carr M.A., Demetres C.A., Demetres C.C., Osburn B.I.;
RT "Association of Bluetongue Virus Gene Segment 5 with Neu"
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases
DR EMBL; U03285; AAA03474.1; -.
DR GO; GO:0019028; C.viral capsid; IFA.
DR GO; GO:0005198; F.structural molecule activity; IFA.
DR InterPro; IPR000145; Orbl_VP5.
DR Pfam; PF00901; Orbl_VP5; 1.
SQ SEQUENCE 526 AA; 59273 MW; D46136DDA1174586 CRC64;

Alignment Scores:
Pred. No.: 140 Length: 526
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Misses: 0
Query Match: 1.24% Indels: 0
DB: 12 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q65756 (1-526)
QY 98 CTATGCATTAGACGACCTTTTCC 75
DB 435 LeuSerLeuSerLeuSerLeuSer 442

RESULT 26
09Y0X9 PRELIMINARY; PRT; 534 AA.
ID 09Y0X9
AC 09Y0X9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Coes54.23.
DN Coes54.23.
OS Ranaid herpesvirus 1 (Lucke tumor herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae.
OC NCBI_TaxID=85655;
RN [1]
RP SEQUENCE FROM N.A.
RA Davidson A.J., Sauerbier W., Dolan A., Addison C., McKinnon R.T.
RT "Genomic studies of the Lucke tumor herpesvirus (RaHV-1) with Neu"
RL EMBL; AF110004; AAD1286.1; -.
DR InterPro; IPR003006; IG_MHC.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 534 AA; 59141 MW; 0873A7A4AC0ACECP9 CRC64;

Alignment Scores:
Pred. No.: 140 Length: 534

```

Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 12 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q9YQX9 (1-534)

QY 1092 CGTGCAGCGTGGAGCCGCGCT 1115
 DB 367 ArgAlaglinArgAlaAlaValAla 374

RESULT 27

ID Q82GP4 PRELIMINARY; PRT; 544 AA.
 AC Q82GP4;
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
 DE Putative integral membrane transport protein.
 GN SAV3853.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=2147403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hamamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osomoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=2147403; PubMed=11572948;
 RA Ikeda H., Ishikawa J., Hamamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.,
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis."
 RL Nat. Biotechnol. 21:526-531(2003).
 DR EMBL; AP005036; BAC71555.1; -.
 KW Complete proteome.
 SQ SEQUENCE 544 AA; 56291 MW; FC13A5045264A3DA CRC64;

Alignment Scores:

Pred. No.: 139 Length: 544
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 16 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q82GP4 (1-544)

QY 1648 ACCTTACCCCTCCCTGACGAC 1671

DB 13 ThnSerThProSerLeuThrSer 20

RESULT 28

ID Q92ZH5 PRELIMINARY; PRT; 545 AA.
 AC Q92ZH5;
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DE Putative arylsulfatase.
 GN RA0510 OR SMA0943.
 OS Rhizobium meliloti (Sinorhizobium meliloti).

OG Plasmid pSymA (megaplasmid 1).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396509; PubMed=11481432;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
 RA Barlow-Hubler F., Bowser L., Capela D., Galibert F., Gouy J.,
 RA Gurjal M., Hong A., Hutzler L., Hyman R.W., Kahn D., Kahn M.L.,
 RA Kaiman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 RA Yen K.-C., Davis R.W., Federspiel N.A., Long S.R.,
 RT "Nucleotide sequence and predicted functions of the entire
 RT Sinorhizobium meliloti pSymA megaplasmid."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 DR EMBL; AF007242; AKK65168.1; -.
 DR PIR; P95325; P95325.
 DR GO; GO:0046821; C:cytochromosomal DNA; IEA.
 DR GO; GO:0008484; F:sulfuric ester hydrolase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR000917; Sulfatase.
 DR Pfam; PF00884; Sulfatase.
 KW plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 545 AA; 60458 MW; 629CC3D9A07A3C13 CRC64;

Alignment Scores:

Pred. No.: 139 Length: 545
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.24% Indels: 0
 DB: 16 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q92ZH5 (1-545)

QY 1593 CGTGAAGTCTGCTGCTGCTGCGA 1570

DB 15 ArgGluValLeuValAlaGlyGly 22

RESULT 29

ID O55887 PRELIMINARY; PRT; 555 AA.
 AC O55887;
 DT 01-JUN-1998 (T-EMBLrel. 06, Created)
 DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE Fusion glycoprotein (F glycoprotein).
 GN F.
 OS Human parainfluenza virus 1.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Respirovirus.
 OX NCBI_TaxID=12730;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=PIV1/Washington/20993/1964;
 RX MEDLINE=75188544; PubMed=166929;
 RA Murphy B.R., Richman D.D., Chaihub E.G., Unlendorf C.P., Baron S.,
 RA Chanock R.M.,
 RT "Failure of attenuated temperature-sensitive influenza A (H3N2) virus
 RT to induce heterologous interference in humans to parainfluenza type 1
 RT virus."
 RL Infect. Immun. 12:62-68(1975).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PIV1/Washington/20993/1964;
 RA Tao T., Durbin A.P., Whitehead S.S., Davoodi F., Collins P.L.,
 RA Murphy B.R.,
 RT "Recovery of a chimeric human parainfluenza virus from PIV3 cDNA in
 RT which the hemagglutinin and fusion glycoproteins of PIV3 were
 RT substituted for those of PIV3."
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN=Washington 1964;
RA Newman S.R., Riggs J.M., Hansen C.T., Collins P.L.,
RM Murphy B.R., Skidopoulos M.H.;
RT "Sequence Analysis of the Washington/1964 Strain of Human
Parainfluenza Virus Type 1 (HPIV1) and Recovery and Characterization
of Wild Type Recombinant HPIV1 Produced by Reverse Genetics.",
RL Virus Genes 24:0-0(2002).
RN (4)
RP SEQUENCE FROM N.A.
RC STRAIN=Washington 1964;
RA Newman S.R., Riggs J.M., Hansen C.T., Collins P.L.,
RM Murphy B.R., Skidopoulos M.H.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF016279; AAC23945.1; -
DR EMBL; AF457102; AAL89407.1; -
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro: IPR000776; Fusion.gly.
DR Pfam: PF00523; fusion.gly.1.
SQ SEQUENCE 555 AA; 60767 MW; 439F9BC2FE1D4EEF CRC64;

Alignment Scores:
Pred. No.: 139 Length: 555
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 12 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x 055887 (1-555)
QY 1193 AGCTTATGCTGCTTCAGCCTG 1170
Db 13 SerLeuLeuSerSerSerSerLeu 20

RESULT 30
ID 055888 PRELIMINARY; PRT; 555 AA.
AC 055888;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Fusion glycoprotein.
GN F.
OS Recombinant PIV3/PIV1 virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae.
OC NCBI_TaxID=65405;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=PIV3-1;
RX MEDLINE=97428580; PubMed=9281512;
RA Durbin A.P., Hall S.L., Siew J.W., Whitehead S.S., Collins P.L.,
RM Murphy B.R.;
RT "Recovery of infectious human parainfluenza virus type 3 from cDNA.",
RL Virology 235:323-332(1997).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=PIV3-1;
RX MEDLINE=98184528; PubMed=9525616;
RA Tao T., Durbin A.P., Whitehead S.S., Davoodi F., Collins P.L.,
RM Murphy B.R.;
RT "Recovery of a fully viable chimeric human parainfluenza virus (PIV)
type 3 in which the hemagglutinin-neuraminidase and fusion
glycoproteins have been replaced by those of PIV type 1.",
RL J.Virol. 72:2955-2961(1998).
DR EMBL; AF016281; AAC23947.1; -
DR GO; GO:0019039; F:Viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro: IPR000776; Fusion.gly.
DR Pfam: PF00523; fusion.gly.1.
SQ SEQUENCE 555 AA; 60768 MW; 439834DB8055CAEF CRC64;

Alignment Scores:
Pred. No.: 139 Length: 555
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 12 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x 055888 (1-555)
QY 1193 AGCTTATGCTGCTTCAGCCTG 1170
Db 13 SerLeuLeuSerSerSerSerLeu 20

RESULT 31
ID 08AAB4 PRELIMINARY; PRT; 559 AA.
AC 08AAB4;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Asparagine synthetase B.
GN BT0551.
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OC NCBI_TaxID=818;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=1263928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.",
RL Science 299:2074-2076(2003).
DR EMBL; AE016928; AAO75658.1; -
DR GO; GO:0004066; P:asparagine synthase (glutamine-hydrolyzing) . . .; IEA.
DR GO; GO:0006529; P:asparagine biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro: IPR001962; Asn_synthase.
DR InterPro: IPR006426; Asn_synth_AEB.
DR InterPro: IPR000583; GATase_2.
DR Pfam: PF00733; Asn_synthase; 1.
DR Pfam: PF00310; GATase_2; 1.
DR TIGRPFam: TIGR01536; asn_synth_AEB; 1.
KM Complete proteome.
SQ SEQUENCE 559 AA; 63019 MW; 292F181736AD4B05 CRC64;

Alignment Scores:
Pred. No.: 139 Length: 559
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 16 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x 08AAB4 (1-559)
QY 1081 AGCGGCCGAGCGTCCAGCGTG 1104
Db 514 SerValProSerValProSerVal 521

RESULT 32
ID 07MG72 PRELIMINARY; PRT; 569 AA.
AC 07MG72;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative oxidoreductase.
GN BB4047.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RB50 / ATCC BAA-568;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parikhili J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Cherevach I.,
 RA Achman M., Atkin R., Baker S., Basham D., Bason N., Doggett J.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Godle A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX64049; CAE34410.1; -.
 KW Complete proteome.
 SQ SEQUENCE 569 AA; 61713 MW; D4F4235C8D86A5BE CRC64;

Alignment Scores:
 Pred. No.: 138 Length: 569
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.24% Indels: 0
 DB: 16 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q7W672 (1-569)

QY 1022 GGACTGAGGGCGTCTCGCTTTTC 999
 Db 241 GlyleuarglyValserleuphe 248

RESULT 33

ID Q7W4P6 PRELIMINARY; PRT; 569 AA.
 AC Q7W4P6;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Purative oxidoreductase.
 GN BP3612
 OS *Bordetella parapertussis*.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12822 / ATCC BAA-587;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parikhili J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Cherevach I.,
 RA Achman M., Atkin R., Baker S., Basham D., Bason N., Doggett J.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Godle A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX64043; CAE38896.1; -.
 KW Complete proteome.
 SQ SEQUENCE 569 AA; 61714 MW; D3D9B5B5BF2EA0DC CRC64;

Alignment Scores:

Pred. No.: 138 Length: 569
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.24% Indels: 0
 DB: 16 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q7W4P6 (1-569)

QY 1022 GGACTGAGGGCGTCTCGCTTTTC 999
 Db 241 GlyleuarglyValserleuphe 248

RESULT 34

ID Q7U0D7 PRELIMINARY; PRT; 573 AA.
 AC Q7U0D7;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Periplasmic glucans biosynthesis protein Mdog.
 GN MDOG OR RB6383.
 OS *Rhodopirellula ballica*.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Pirellula.
 OX NCBI_TaxID=117;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1;
 RX MEDLINE=22735913; PubMed=12835416;
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Borzym K., Hettmann K., Rabus R.,
 RA Schlesner H., Amann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete *Pirellula* sp.
 RT strain 1.";
 RT Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
 DR EMBL; BX294144; CAD74766.1; -.
 KW Complete proteome.
 SQ SEQUENCE 573 AA; 63889 MW; E8E5381E34A95DEF CRC64;

Alignment Scores:
 Pred. No.: 138 Length: 573
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.24% Indels: 0
 DB: 16 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q7U0D7 (1-573)

QY 59 AGAGCCAGCAAGCTCAGTAACA 36
 Db 457 ArgAlaThrAsnLeuSerValThr 464

RESULT 35

ID Q9P8M8 PRELIMINARY; PRT; 605 AA.
 AC Q9P8M8;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Glucose oxidase.
 GN GO2.
 OS *Aspergillus niger*.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; *Aspergillus*.
 OX NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ACMO4;
 RA Kim M., Kwon T., Yang M., Kim D.;
 RT "Cloning and identification of a unique allele of the glucose oxidase
 RT from a Korean wild type strain of *Aspergillus niger*.";

Query Match:	1.24%	Indels:	0
DB:	4	Gaps:	0
US-10-029-345A-108_COPY_532_2532 (1-2001) x OGN4M1 (1-677)			
Qy	1493	TSAGGCGGTCACATCGGTCGGT	1470
Db	588	TPPALVALPProleuLeuVal	595
RESULT 40			
065201		PRELIMINARY;	PRT; 692 AA.
ID	065201		
AC	065201;		
DT	01-AUG-1998 (TREMBLrel_07, Created)		
DT	01-AUG-1998 (TREMBLrel_07, Last sequence update)		
DE	01-JUN-2003 (TREMBLrel_24, Last annotation update)		
GN	Acyl-CoA oxidase (EC 1.3.3.6).		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids		
OC	euroside II; Brassicales; Brassicaceae; Arabidopsiis.		
OX	NCBI_TaxId=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV. Columbia; TISSUE=Seedling hypocotyl;		
RA	MEDLINE=20040045; PubMed=10571860;		
RT	Hooks M.A., Kellis F., Graham I.A.;		
RL	"Long-chain acyl-CoA oxidases of Arabidopsis.",		
DR	Plant J. 20:1-13(1999).		
DR	EMBL; AF057043; AAC13497.1; -.		
DR	PIR; T52120; T52120.		
DR	GO; GO:0005777; C:peroxisome; IEA.		
DR	GO; GO:0003997; F:acyl-CoA oxidase activity; IEA.		
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.		
DR	GO; GO:0006118; P:electron transport; IEA.		
DR	GO; GO:0006635; P:fatty acid beta-oxidation; IEA.		
DR	InterPro; IPR002655; AC0X.		
DR	InterPro; IPR006090; Acyl-CoA_dh_C.		
DR	Pfam; PF01756; AC0X; 1.		
DR	Pfam; PF00441; Acyl-CoA_dh; 1.		
KW	Oxidoreductase.		
SQ	SEQUENCE 692 AA; 77496 MW; 59D134E334772D24 CRC64;		
Alignment Scores:			
Pred. No.:	135	Length:	692
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	10	Gaps:	0
US-10-029-345A-108_COPY_532_2532 (1-2001) x 065201 (1-692)			
Qy	596	AGCGTACCTTATCCCGACGCTC	619
Db	34	SetLeuThLeuSerProSerLeu	41
RESULT 41			
08ZK16		PRELIMINARY;	PRT; 694 AA.
ID	08ZK16		
AC	08ZK16;		
DT	01-MAR-2002 (TREMBLrel_20, Created)		
DT	01-MAR-2002 (TREMBLrel_20, Last sequence update)		
DT	01-JUN-2003 (TREMBLrel_24, Last annotation update)		
DE	Putative ATP-dependent Ion protease.		
GN	STM4491.		
OS	Salmonella typhimurium.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
OC	Enterobacteriaceae; Salmonella.		
OX	NCBI_TaxId=602;		
RN	[1]		
RP	SEQUENCE FROM N.A.		

RC STRAIN=LT2 / SCSCL412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Potwilk S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 DR EMBL: AEO08911; AL23309.1; -;
 DR GO: GO:0008233; F:Peptidase activity; IEA.
 KM Protease; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 694 AA; 77463 MW; 53BE7B95B5F8B878 CRC64;

Alignment Scores:
 Pred. No.: 135 Length: 694
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 16 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q8ZK16 (1-694)

QY 970 CCTGTCCTGCTGCTCAGAGGT 993
 DB 8 ProValProAlaValSerGluGly 15

RESULT 42
 Q8U1R9 PRELIMINARY; PRT; 707 AA.
 ID Q8U1R9
 AC Q8U1R9;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein PF1136.
 GN PF1136.
 OS *Pyrococcus furiosus*.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Wise R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the *Pyrococcus furiosus* genome.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AEO10223; AL81260.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 707 AA; 80090 MW; 012ACBC9C987A6D7 CRC64;

Alignment Scores:
 Pred. No.: 134 Length: 707
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 17 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q8U1R9 (1-707)

QY 67 AGTGAACGGAAGAGTCTGCTA 90
 DB 53 SerGlyThrGluysValLeu 60

RESULT 43
 P87583 PRELIMINARY; PRT; 711 AA.
 ID P87583
 AC P87583;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Non-structural protein 1.
 GN NS1.
 OS Chikungunya virus.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=56820;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99119498; PubMed=9918883;
 RA Yoo B.C., Lee D.H., Park S.M., Park J.W., Kim C.Y., Lee H.S.,
 RA Seo J.S., Park K.J., Ryu W.S.;
 RT "A novel parvovirus isolated from Manchurian chipmunks.";
 RL Virology 253:250-258(1999).
 DR EMBL: U86868; AA882733.1; -;
 DR InterPro: IPR001257; Parvo_NSI.
 SQ SEQUENCE 711 AA; 78578 MW; 75E64EB8563BAEB CRC64;

Alignment Scores:
 Pred. No.: 134 Length: 711
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.24% Indels: 0
 DB: 12 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x P87583 (1-711)

QY 1647 CTGGGGGGCCAGATATCCAGTG 1624
 DB 404 LeuGlyGlyGlnAspIleArgVal 411

RESULT 44
 Q8Y2P4 PRELIMINARY; PRT; 727 AA.
 ID Q8Y2P4
 AC Q8Y2P4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein A114378.
 GN A114378.
 OS *Anabaena* sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Matanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium *Anabaena* sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL: AP003596; BAB76077.1; -;
 DR PIR: AB2353; AB2353.
 DR InterPro: IPR008941; TPR-1like.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 727 AA; 81109 MW; 78E7D9B9FA603183 CRC64;

Alignment Scores:
 Pred. No.: 134 Length: 727
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 16 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q8Y2P4 (1-727)

QY 1081 AGCGTGCCAGCGTCCAGCGTG 1104
 DB 116 SerValProSerValProSerVal 123

```

RESULT 45
Q98S05
ID 098S05 PRELIMINARY; PRT; 752 AA.
AC 098S05;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, last sequence update)
DE Cell division cycle protein 48 homolog.
GN CDC48.
OS Giardia theta (Cryptomonas phl.).
OC Nucleomorph.
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guiliardia.
ON NCBI_TaxID=55529;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21223349; PubMed=11323671;
RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
RA Wu X., Reith M., Cavalier-Smith T., Mäler U.G.;
RT "The highly reduced genome of an enslaved algal nucleus.";
RL Nature 410:1091-1096(2001).
DR EMBL; AF083031; AK39773.1; -.
DR PIR; H90135; H90135.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0008151; P:cell growth and/or maintenance; IEA.
DR InterPro; IPR003959; AAA_ATPase.
DR InterPro; IPR005938; AAA_CDC48.
DR InterPro; IPR003860; AAA_sub.
DR InterPro; IPR009010; Asp_decarb_fold.
DR InterPro; IPR00338; ATPaseVar_N.
DR Pfam; PF00004; AAA; 2.
DR SMART; SM00382; CDC48_N; 1.
DR TIGRPFAMs; TIGR01243; CDC48; 1.
DR PROSITE; PS00674; AAA; 2.
KW ATP-binding.
SQ SEQUENCE 752 AA; 84454 MW; FE59364D5A82BD71 CRC64;

Alignment Scores:
Pred. No.: 133
Score: 8.00 Length: 752
Percent Similarity: 100.00% Matches: 8
Best Local Similarity: 100.00% Conservative: 0
Query Match: 1.24% Mismatches: 0
Indels: 0
Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q98S05 (1-752)
QY 771 AGCGATGCTGCGGAGCGGAGAT 748
DB 570 SerAspGlyGlyGlyAlaGlyAsp 577

RESULT 46
Q9KIB0
ID 09KIB0 PRELIMINARY; PRT; 757 AA.
AC 09KIB0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DE Hypothetical tonB-linked outer membrane receptor Pg13.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
ON NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=WS0.
RA Ross B., Barr I., Patterson M., Agius C., Rothel L., Margetts M.,
RA Hocking D., Webb E.;
RT "P. gingivalis polypeptides and nucleic acids.";

```

```

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WS0;
RA Ross B.C.;
RT "Genomic analysis of Porphyromonas gingivalis for vaccine discovery.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE (BY SIMILARITY).
DR EMBL; AF237558; AAF81416.1; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; F:transport; IEA.
DR InterPro; IPR008969; Carboxypeptid_reg.
DR InterPro; IPR000531; TonB_box.
DR Pfam; PF00593; TonB_dep_Rec; 1.
KW Membrane; Outer membrane; Receptor; TonB box.
SQ SEQUENCE 757 AA; 84517 MW; 224B6D5264F9D62 CRC64;

Alignment Scores:
Pred. No.: 133
Score: 8.00 Length: 757
Percent Similarity: 100.00% Matches: 8
Best Local Similarity: 100.00% Conservative: 0
Query Match: 1.24% Mismatches: 0
Indels: 0
Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q9KIB0 (1-757)
QY 1187 TTGCTGCTTCAGCGCTGCTGCG 1164
DB 16 LeuLeuSerSerSerLeuSerAla 23

RESULT 47
Q91CZ2
ID 091CZ2 PRELIMINARY; PRT; 759 AA.
AC 091CZ2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; Circoviridae; Anellovirus.
ON NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21488921; PubMed=11601907;
RA Okamoto H., Nishizawa T., Takahashi M., Asabe S., Tsuda F.,
RA Yoshikawa A.;
RT "Heterogeneous distribution of TT virus of distinct genotypes in
RT multiple tissues from infected humans.";
RL Virology 288:358-368(2001).
DR EMBL; AB060595; BAB69912.1; -.
DR InterPro; IPR004219; TTvirus_Unk.
DR Pfam; PF02956; TT_ORF1; 1.
SQ SEQUENCE 759 AA; 89860 MW; 5B79DFD71A37010D CRC64;

Alignment Scores:
Pred. No.: 133
Score: 8.00 Length: 759
Percent Similarity: 100.00% Matches: 8
Best Local Similarity: 100.00% Conservative: 0
Query Match: 1.20% Mismatches: 0
Indels: 0
Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q91CZ2 (1-759)
QY 389 TCTCTGTTGTTTCCTGCGCTCT 412
DB 249 SerLeuValValSerLeuAlaSer 256

RESULT 48
Q8L3X5

```

ID 08L3X5 PRELIMINARY; PRT; 773 AA.
 AC 08L3X5;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative reverse transcriptase.
 GN OSJNA0036D19.11 OR OSJNA0050E08.4.
 OS Oryza sativa (Rice), and
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=4530, 39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O.sativa, and O.sativa (japonica cultivar-group);
 RA Wang R.A., Yu Y., Yang T.J., Nah G., Soderlund C., Chen M., Kim H.-R.,
 RA Rambo T., Sakki C., Henry D., Oates R., Simmons J.;
 RT "Rice Genomic Sequence";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O.sativa (japonica cultivar-group); STRAIN=cv. Nipponbare;
 RA The Rice Chromosome 10 Sequencing Consortium;
 RT "In-depth view of structure, activity, and evolution of rice
 chromosome 10";
 RL Science 300:1566-1569(2003).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O.sativa (japonica cultivar-group); STRAIN=cv. Nipponbare;
 RA Buehl C.R., Ming R.A., McCombie W.R., Messing J., Yuan Q.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC116600; AA004143.1; -
 DR EMBL; AC098696; AA004203.1; -
 DR EMBL; AB017059; AAP52227.1; -
 DR Gramene; Q8L3X5; -
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 KM RNA-directed DNA polymerase
 SQ SEQUENCE 773 AA; 85310 MW; DB5DDC866CC086A3 CRC64;

Alignment Scores:
 Pred. No.: 133 Length: 773
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.24% Indels: 0
 DB: 10 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x Q8L3X5 (1-773)
 QY 1957 TGGCCACTTTCCTCCAGCTCTTCCC 1934
 DB 303 CyaproleuserProalaleupro 310

RESULT 49
 O8UGX6 PRELIMINARY; PRT; 789 AA.
 ID O8UGX6;
 AC O8UGX6;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Na+/H+ antiporter.
 GN NMHA OR ATU0909 OR AGR_C_1658.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 OC NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Boyce D. Sr.,

RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayvin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung W., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Neiser E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Quirillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houtmel K., Gordon J., Vaudin M., Iarchenko O., Epp A., Liu F.,
 RA Mollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009056; AA141923.1; -
 DR EMBL; AE008022; AA86713.1; -
 DR PIR; AE2688; AE2688.
 DR PIR; H97469; H97469.
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
 DR InterPro; IPR003918; NADHb oxidred4.
 DR InterPro; IPR003916; NADHb oxidred5.
 DR InterPro; IPR001750; Oxidored_q1.
 DR InterPro; IPR001516; Oxidored_q1_N.
 DR Pfam; PF00361; Oxidored_q1_1.
 DR Pfam; PF00662; Oxidored_q1_N_1.
 DR PRINTS; PR01434; NADHGNASS5.
 DR PRINTS; PR01437; NUOXDRPASE4.
 KM Complete proteome.
 SQ SEQUENCE 789 AA; 84323 MW; 0329652251F2D5FB CRC64;

Alignment Scores:
 Pred. No.: 132 Length: 789
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.24% Indels: 0
 DB: 16 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x O8UGX6 (1-789)
 QY 1869 CTTTCAAGGGGCTCTTCATG 1846
 DB 340 LeuphelysclYAlaleuphewet 347

RESULT 50
 O92RA3 PRELIMINARY; PRT; 791 AA.
 ID O92RA3;
 AC O92RA3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative cation efflux system protein.
 GN PHA2 OR R00992 OR SMC00051.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OC NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RA MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kies E., Lelaure V., Masny D.,

RA Pohl T., Portetelle D., Puehler A., Purnelle B., Rampeger U.,
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT *Sinorhizobium meliloti* strain 1021.",
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL; AL591785; CAC45564.1; -.
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
 DR InterPro; IPR003918; NADHdb_oxred4.
 DR InterPro; IPR003916; NADHdb_oxreds.
 DR InterPro; IPR001750; Oxidored_q1.
 DR InterPro; IPR001516; Oxidored_q1.
 DR Pfam; PF00361; oxidored_q1; 1.
 DR Pfam; PF00662; oxidored_q1_N; 1.
 DR PRINTS; PR01434; NADHDGNASE5.
 DR PRINTS; PR01437; NUOXDRDTASE4.
 KW Complete proteome.
 SQ SEQUENCE 791 AA; 84181 MW; 7D833BD253BD9FF3 CRC64;

Alignment Scores:

Pred. NO.:	132	Length:	791
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.24%	Indels:	0
DB:	16	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x Q92RA3 (1-791)

QY 1869 CTTTCAAGGGGCTCTTCATG 1846

DB 338 LeuphelysGlyAlaLeuphemeT 345

Search completed: June 21, 2004, 13:13:30
 Job time : 148.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 12:46:21 ; Search time 94.5 Seconds

(without alignments)
11965.663 Million cell updates/sec

Title: US-10-029-345a-108_COPY_532_2532

Sequence: 1 gttgtcatcgtccatgagat.....tggaaatcatcgtggtctcc 2001

Scoring table:

OLIGO	
Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Word size: 1

Total number of hits satisfying chosen parameters: 324566

Minimum DB seq length: 473

Maximum DB seq length: 200000000

Post-processing: listing first 75 summaries

Command line parameters:

-MODE=frame+np.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool_p/US10029345/runat_21062004_124614_5672/app_query.fasta_1.2183
-DB=A_Geneseq_29345 -QPMT=faatan -SUFFIX=oligo.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=75 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL
-OUTPM=pro -NORM=ext -HEAPSIZE=500 -MINLEN=473 -MAXLEN=200000000
-USR=US10029345_@CGN_1.1_81@runat_21062004_124614_5672 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

A_Geneseq_29345:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	665	99.7	665	5	ABR52381 Protein r
2	665	99.7	665	5	ABR52407 Protein r
3	472	70.8	517	5	Aau79159 Human dua
4	472	70.8	665	4	AAE04834 Human SGP
5	472	70.8	665	4	AAU09016 Human dua
6	472	70.8	665	5	AAU79136 Human dua
7	472	70.8	665	5	AAU09946 Protein s
8	472	70.8	665	5	AAU75789 Human pro
9	472	70.8	665	5	ABR97946 Human pro
10	472	70.8	665	5	AAU79929 Human dua

11	472	70.8	665	5	ABR97291	AbB97291 Novel hum
12	472	70.8	665	5	ADA54744	AdA54744 Human pro
13	452	67.8	665	5	Aau79161	Aau79161 Human dua
14	421	63.1	665	5	AAU79162	Aau79162 Human dua
15	350	52.5	665	5	ABR52352	AbR52352 Protein r
16	318	47.7	666	4	AAAB20335	AAb20335 Human pro
17	281	42.1	664	4	ABR52424	AbR52424 Protein r
18	258	38.7	672	4	AAW52744	AAw52744 Human pro
19	64	9.6	660	5	ABR52385	AbR52385 Protein r
20	20	3.0	579	7	ADB08458	AdB08458 Novel pro
21	20	3.0	625	5	ABR52382	AbR52382 Protein r
22	20	3.0	625	5	ABR52350	AbR52350 Protein r
23	20	3.0	625	6	ABG73440	ABg73440 Human dua
24	20	3.0	663	2	AAW29150	AAw29150 Dual-spec
25	20	3.0	663	5	ABR52351	AbR52351 Protein r
26	13	1.9	476	4	ABR63527	ABr63527 Drosophila
27	12	1.8	836	4	ABG00724	ABg00724 Novel hum
28	10	1.6	498	6	ABU29344	ABu29344 Protein e
29	8	1.2	473	7	ADA44807	AdA44807 CD4/TCR C
30	8	1.2	474	3	AAV59170	AAv59170 CD4-Ig fu
31	8	1.2	481	1	AAV93011	AAp93011 Genetic C
32	8	1.2	481	3	AAV19510	AAv19510 CD4-Igm E
33	8	1.2	481	3	AAV51081	AAv51081 Human fus
34	8	1.2	481	3	AAV59171	AAv59171 CD4-Ig fu
35	8	1.2	502	7	ADB08675	AdB08675 Novel pro
36	8	1.2	509	2	AAAB00158	AAa00158 eCD4-SCFV
37	8	1.2	509	2	AAW26792	AAw26792 Mouse TIE
38	8	1.2	509	2	AAW47533	AAw47533 Amino ac1
39	8	1.2	509	2	AAV23734	AAv23734 TIE ligand
40	8	1.2	509	3	AAV90398	AAv90398 Mouse TIE
41	8	1.2	509	7	ADD69365	AdD69365 Human Ang
42	8	1.2	519	2	AAAR20152	AAa20152 Human CD4
43	8	1.2	524	1	AAV94703	AAp94703 Sequence
44	8	1.2	525	3	AAAG51289	AAg51289 Arabidops
45	8	1.2	525	3	AAAG22387	AAg22387 Arabidops
46	8	1.2	530	2	AAAR6783	AAr6783 CD4-IgG2
47	8	1.2	530	2	AAAR6679	AAr6679 CD4-IgG2
48	8	1.2	530	3	AAAY85080	AAy85080 CD4-IgG2
49	8	1.2	530	4	AAAB67323	AAb67323 CD4-IgG2
50	8	1.2	530	4	AAAB80884	AAb80884 Human CD4
51	8	1.2	530	6	ABG71123	ABg71123 CD4-Immun
52	8	1.2	532	2	AAAR7278	AAr7278 CD4:gamma
53	8	1.2	532	2	AAAR8678	AAr8678 T-cell re
54	8	1.2	532	2	AAAR89458	AAr89458 CD4:eta f
55	8	1.2	532	2	AAW02215	AAw02215 CD4:T-cel
56	8	1.2	532	2	AAW83141	AAw83141 Chimeric
57	8	1.2	534	2	AAAR6531	AAr6531 Sequence
58	8	1.2	534	5	AAAB71495	AAb71495 P. amygd
59	8	1.2	540	2	AAW70460	AAw70460 South Afr
60	8	1.2	540	2	AAW70462	AAw70462 Girwood
61	8	1.2	540	2	AAW70464	AAw70464 Sindbis v
62	8	1.2	549	2	AAAR04920	AAa04920 Immunopro
63	8	1.2	557	2	AAAR04923	AAa04923 Immunopro
64	8	1.2	557	2	AAAR04919	AAa04919 Immunopro
65	8	1.2	575	2	AAAR27276	AAr27276 CD4:zeta
66	8	1.2	575	2	AAAR78676	AAr78676 T-cell re
67	8	1.2	575	2	AAAR89456	AAr89456 CD4:zeta
68	8	1.2	575	2	AAW02213	AAw02213 CD4:T-cel
69	8	1.2	575	2	AAW83140	AAw83140 Chimeric
70	8	1.2	576	5	AAE21625	AAe21625 Human gen
71	8	1.2	577	2	AAAR04924	AAa04924 Immunopro
72	8	1.2	584	4	AAAB98948	AAb98948 Novel hum
73	8	1.2	585	4	AAAB98947	AAb98947 Novel hum
74	8	1.2	590	6	ABU07697	ABu07697 Viral coa
75	8	1.2	605	2	AAAR04724	AAa04724 Glucose o

ALIGNMENTS

RESULT 1
ABR52381
ID ABR52381 standard; protein; 665 AA.
XX

AC ABR52381;
XX
DT 19-JUN-2003 (first entry)
DE
XX Protein relating to the invention SEQ ID NO: 109.
DE
XX
XX antiproliferative; hepatotropic; nephrotropic; antiarthritic;
KW antiproliferative; cardiant; cytosstatic; gene therapy; liver disease;
KW proliferative disorder; renal failure; cardiovascular disorder;
KW immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
OS Homo sapiens.
XX
XX MO200257460-A2.
XX
XX 25-JUL-2002.
XX
XX 20-DEC-2001; 2001WO-US050459.
XX
XX 20-DEC-2000; 2000US-0256686P.
XX 03-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Jackson DG, Feder J, Nelson T, Mintier G, Ramathanan C, Lee L,
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
PI Krysek S, Mattee P, Suchard S, Bamas D;
XX
XX WPI: 2002-599721/64.
DR N-PSDB; ACC60559.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
XX disorders.
XX
XX
XX Claim 5; Fig 12; 801pp; English.

CC The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antiarthritic, antiprosstatic, cardiant, and cytosstatic activity. The
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention

XX
XX Sequence 665 AA;
XX
XX

Alignment Scores:
Pred. No.: 0
Score: 665
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 99.70%
DB: 5

Length: 665
Matches: 665
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x ABR52381 (1-665)

7 AAGGCCCATGAATGATTGGAATCTCAATTGTTACTGAGAGTTGGTGGCTGCTGGAA 66
db 1 MetAlaHisLeuIleCilleglThrcInlIleValThrcIuArgLeuValAlaIleuLeuGlu 20
67 AGTGAACGGAANAAGTCTCTAATTGATGATGCCGCCGCAATTGTGGAAATACAAATCATCC 126

[illegible]

Db 381 SerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAlaLeuGluValArgSerPhe 400
 QY 1207 TCTCTGGATATCAATTCAGTTTATTCAGCCAGCATGGCAGCATCTTACATGGCTTC 1266
 Db 401 SerLeuAspIleLeuSerValSerIYrSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 QY 1267 TCCCTCATCAGAAAGATGCTTGGAAATCTACAACCTTCCACTCTCTGATGGAGCAAC 1326
 Db 421 SerSerSerGluAlaAlaLeuGluIYrYrLeuProSerThrThrLeuAspGlyThrAsn 440
 QY 1327 AACCTATGCCAGTTCTCCCTGTTCCAGAACTATCGAGCAGACTCCCGAAGCACTGCT 1386
 Db 441 LysLeuGlySerGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro 460
 QY 1387 GATTAAGAGGAGACAGCAGCATCTCCCAAGAGCTGACAGCCGCGCTTACAGACCCAG 1446
 Db 461 AspLysGluGlnIuAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
 QY 1447 AGCAGAGGATTTGATTCGGTCCAGACAGCAGAGTGGACCGCGCCAGAGTCCCTTTTA 1506
 Db 481 SerLysArgLeuHisSerValArgIuSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
 QY 1507 TCTCCACTGCATGAAAGTGAGAGCCTGAGAGACATTTACCAACAGCTTCTTTTCGGC 1566
 Db 501 SerProLeuHisArgSerGlySerValGluAspAsnIYrHisThrSerPheLeuPheGly 520
 QY 1567 CTTTCCACAGCCAGCAGCAGCCTTCAGAGAGTCTGCTGGCTGGGCTTAAAGGCTGGCAC 1626
 Db 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheHis 540
 QY 1627 TCGGATATCTTTGGCCCCCGCAGACCTCAACCCCTTCCCTGACAGCAGCTGATTTTGGC 1686
 Db 541 SerAspIleLeuAlaProGlnIuSerThrProSerLeuThrSerSerIYrPheAla 560
 QY 1687 ACAGAGTCTCTCACACTTCTACTCTGCTCAGCCATCTACCGAGAGCAGTGGCACTTACTCT 1746
 Db 561 ThrGluSerSerHisIlePheIYrSerAlaSerAlaIleIYrGlyGlySerAlaSerIYrSer 580
 QY 1747 GCTTACAGCTGACGCCAGCTGCCCACTTGGGAGACCAAGCTTATTTCTGTGGCAGCGG 1806
 Db 581 AlaIYrSerCysSerGlnLeuProIuIYrGlyAspGlnValIYrSerValArgArgArg 600
 QY 1807 CAGAACCGAAGTACAGAGCTGACTCGCGCGAGAGCTGGAGTAAAGAGACCCCTTTGAA 1866
 Db 601 GlnLysProSerAspArgAlaAspSerArgArgSerIYrIleGluLeuSerProPheGlu 620
 QY 1867 AAGCAGTTTAAACGACAGAGCTCCAAATGAATTTGAGAGAGCATCATGTACAGAAAC 1926
 Db 621 LysGlnPheLysArgArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn 640
 QY 1927 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCAAGCTTTCGCGGACAGCATGAA 1986
 Db 641 ArgSerArgGlnGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGlu 660
 QY 1987 ATCATGTAGGCTTCC 2001
 Db 661 IleIleGluValSer 665
 RESULT 2
 ABR52407 standard; protein; 665 AA.
 AC ABR52407;
 XX 19-JUN-2003 (first entry)
 DE Protein relating to the invention SEQ ID NO: 148.
 XX anti-proliferative; hepatotropic; nephrotropic; anti-arthritic;
 KM anti-proliferative; cardiatic; cytostatic; gene therapy; liver disease;
 KM proliferative disorder; renal failure; cardiovascular disorder;
 KM immunological disorder; arthritis; psoriasis; congenital heart defect;
 KM congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.

XX OS Homo sapiens.
 XX EN WO200257460-A2.
 XX PD 25-JUL-2002.
 XX PF 20-DEC-2001; 2001WO-US050459.
 XX PR 20-DEC-2000; 2000US-0256868P.
 XX PR 30-MAR-2001; 2001US-0280186P.
 XX PR 01-MAY-2001; 2001US-0287735P.
 XX PR 05-JUN-2001; 2001US-0295648P.
 XX PR 25-JUN-2001; 2001US-0300465P.
 XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX PI Jackson DG, Feder J, Nelson T, Mantler G, Ramanathan C, Lee L;
 PI Slemers N, Bol D, Schieven G, Finger J, Todderud CG, Basolino D;
 PI Kryszek S, Mcatee P, Suchard S, Banas D;
 DR WPI; 2002-599721/64.
 DR N-PSDB; ACC60572.
 PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 XX PS Disclosure; Fig 19; 801pp; English.
 XX CC The invention relates to a novel isolated nucleic acid comprising a
 CC polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotides fully defined in the specification. The polynucleotide of
 CC the invention has anti-proliferative, hepatotropic, nephrotropic,
 CC anti-arthritic, anti-proliferative, cardiatic, and cytostatic activity. The
 CC polynucleotide may have a use in gene therapy. A polynucleotide or
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention
 CC XX
 SQ Sequence 665 AA:
 Alignment Scores:
 Pred. No.: 0 Length: 665
 Score: 665.00 Matches: 665
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.70% Indels: 0
 DB: 5 Gaps: 0
 US-10-029-345A-108_COPY_532_2532 (1-2001) x ABR52407 (1-665)
 QY 7 ATGGCCCATGAGATGATTTGAACTCAATTTGATGAGAGGTGGCTGTGCGAA 66
 Db 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 QY 67 AGTGAACGGAAAAAGTGTCTTAATTTAGTACGGCCGCTTTTGGAAATACATATC 126
 Db 21 SerGlyThrGluLysValLeuLeuLeuAspSerArgProPheValGluIYrAsnThrSer 40
 QY 127 CACATTTGGAAGCCCTTAATATCACTGCTCCAGCTTATGAGGAGGAAGTTGCAACG 186
 Db 41 HisIleLeuGlnIuAlaIleAsnIleLeuGlnCysSerLysLeuMetLysArgArgLeuGlnGln 60
 QY 187 GACAAAGTTTAATTAACAGAGCTCATCGAGTTCAGCGAAACATTAAGTTGACATGAT 246
 Db 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaIYrHisIleLysValAspIleAsp 80

QY	427	TGAGTCACAAAGGTTTACTGTTTACATGCATCAAGCTCCCAAGATGTTGGCTCTCTCTTCA	306
Db	81	CysSerGlnIysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSer	100
QY	307	GACTGTTTCTCACTGTACTTCTGGGTAAACCTGGAGAGACCTTCACTCTGTACCTG	366
Db	101	AspCysPheLeuThrValLeuLeuGlnIySLeuGlnIySserPheAsnSerValHisLeu	120
QY	367	CTTGCAGGCGGGTTTGCTGAGTTCTCTCGTTGTTTCCCTGGCCCTCTGTGAAGAAATCC	426
Db	121	LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlySglnGlyIySser	140
QY	427	ACTTAAGTCCCTACCTGCATTTCTCAGCCCTTGCTTACCTGTGGCAACATTTGGCCAAAC	486
Db	141	ThrLeuValProThrCysAlaSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
QY	487	CGAATCTTCCCAATCTTTATCTTTGGCTGCCAGGAGATGCTCTCAACAGAGCTGATA	546
Db	161	ArgIleLeuProAsnLeuTyrLeuGlnIyCysGlnArgAspValLeuAsnIySglnLeuIle	180
QY	547	CAGCAGAAATGGGATTTGGTTATGTGTAAAGCCAGCTATACCTGTCCAAAGCTGACTTT	606
Db	181	GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyrThrCysProIySProAspPhe	200
QY	607	ATCCCCAGAGTCACTTTCTCGCGTGCCTGTGAATGACAGCTTTTGTGAGAAAATTTTG	666
Db	201	IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIySLeu	220
QY	667	CCGAGTGTGGCAAAATCAGTAGAATTTTCATTGAGAAAGCAAAAGCCCTCCATGATGTGT	726
Db	221	ProThrLeuAspIySserValAspPheIleGlnIyAlaIyAlaSerAsnIyCysVal	240
QY	727	CTAGTCACCTGTTTAAAGTGGGATCTCCCGCTCCGACCATGAGCTATGCGCTCACTCATG	786
Db	241	LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMet	260
QY	787	AAGAGATGAGCAATGTCCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAGAACTTACT	846
Db	261	LysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValIySglnIyAspProThr	280
QY	847	ATATCTCCAACTTCAATTTTCTGGGCAACCTCTGACCTATGAGAAAGAAATTTAAGAC	906
Db	281	IleSerProAsnPheAsnPheLeuGlnIleuLeuAspTyrGlnIySAspIleIyAsn	300
QY	907	CAGACTGGAGCAATCAGGGGCCAAAGAACAACTCAAGCTGTCGACTTGAGAGAACCAAT	966
Db	301	GlnThrGlyAlaSerClyProIySserIySLeuIySLeuHisIleGlnIySProAsn	320
QY	967	GAACTGTCTCCGTGTCTCAGAGGCTGAGCAAAAGCCGAGAGCCCTCAGTCCACC	1026
Db	321	GlnProValProAlaValSerGlnIyGlnIySserGlnIySProLeuSerProPro	340
QY	1027	TGNGCGAATCTGACACTCAGAGGACAGACAGCAAAAGCCCGTCAATCCCGCAGCGTG	1086
Db	341	CysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnIySProValHisProAlaSerVal	360
QY	1087	CCAGCGTGGCCAGCGTCCAGCGCTGCTTAAAGACAGCCCGCGTGAACAGGGCGTC	1146
Db	361	ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu	380
QY	1147	AGTGGGCTGCACCTGTCCGAGACAGAGCTGGAGACAGCATATAGCTCAAGCGTTCCTTC	1206
Db	381	SerGlyLeuHisIleuSerAlaAspArgLeuGlnAspSerAsnIySLeuIySAspSerPhe	400
QY	1207	TCTTGATATCAATCAGTTTCATATTTTCAGCCAGCTGGCAGACATCTTACATGGCTTC	1266
Db	401	SerLeuAspIleIySserValSerTyrSerHisSerMetAlaAlaSerLeuHisGlyPhe	420
QY	1267	TCCATCATCAGAAGATGCTTTGGAAATACTACAAACTTTCCATCATCTGTGATGGACCAAC	1326
Db	421	SerSerSerGlnAspAlaLeuGlnIyTyrIySProSerThrThrLeuAspGlyIyHisAsn	440
QY	1327	AAGCTATGCCAGTTTCTCCCTGTTTCAGGAACTATGAGACAGACTCCGAAACAGTCTCT	1386

DB	Accession	Protein Name	Protein Description	Protein Length
Db	441	LysineCysGlnPheSerProValGlnGlnLysSerGlnGlnThrProGlnThrSerPro		460
QY	1387	GATTAAGAGGAAGCCAGCATCCCCAAGAAGCTGACAGCCGACGCTTACAGACAGCAG		144
Db	461	AspYsGlnGlnValSerIleProGlyLysLeuGlnThrAlaArgProSerAspSerGln		480
QY	1447	AGCAAGGATTCGATTGTTGGTTCAGAACCAAGCAGCAGCAGCCGCCAGAGGTCCCTTTTA		150
Db	481	SerYsArgLeuHisSerValArgThrSerSerSerGlyThrAlaArgSerLeuLeu		500
QY	1507	TCTCCATCGATCGAATGTGGAGCGTGAGAGCAATTACACACCAAGCTTCCCTTTTCGGC		156
Db	501	SerProLeuHisArgSerGlySerValGlnAspAsnThrIleThrSerPheLeuPheGln		520
QY	1567	CTTTCCACCAAGCAGCAGCAGCAGCTCACGAAGTCTGGCTGGGCTTTAAGGCTTGAC		162
Db	521	LeuSerThrSerGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysSerGlyTrpHis		540
QY	1627	TCGGATATCTGGCCCCCAGACACTTACCCTTCCCTTACACAGCAGCTGTATTTGGC		168
Db	541	SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpYrPheAla		560
QY	1687	ACAGAGTCTTCACACTTCTACTCTGGCTGACGACATCTACGAGGAGGAGCCAGTAACTCT		174
Db	561	ThrGlnSerSerHisPheThrSerHisAlaIleTyrGlyGlySerAlaSerTyrSer		580
QY	1747	GCCTAAGCTGACGACGCTGCTGCCACTTGGCGAGACCAAGTCTATCTGTGGCAGGCGG		180
Db	581	AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg		600
QY	1807	CAGAACCCAAAGTACAGAGCTACTCGCGCGGACCTGGCAATGAAAGAACGCCCTTTGAA		186
Db	601	GlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGlnGlnSerProPheGln		620
QY	1867	AAGCAGTTTAAACGAGAAAGCTGCCAAATGGAATTTGGAGAGACATCATGTACAGAAC		192
Db	621	LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn		640
QY	1927	AGGTACACGGGAAGAGCTGGGGAAAGTGGGCGAGTCAAGTCTTTCGGCAGCATGAA		198
Db	641	ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln		660
QY	1987	ATCATTTGAGGTCTCC 2001		
Db	661	IleIleGlnValSer 665		
RESULT 3				
ID	AAU79159	standard; protein; 517 AA.		
AC	AAU79159;			
XX	02-JUL-2002	(first entry)		
DE	Human dual-specificity phosphatase-3 (DSP-16) alternative form protein.			
XX	Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;			
KW	mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;			
KW	cancer; graft-versus-host disease; allergy; metabolic disease;			
KW	abnormal cell growth; abnormal cell proliferation; contact inhibition;			
KW	cell cycle abnormality; anchorage independent cell growth; apoptosis;			
KW	intercellular adhesion; DSP-16 modulator; chromosome 12p.			
XX	Homo sapiens.			
OS	Key	Location/Qualifiers		
XX	Domain	94..103		
XX	/label= Active_site_domain			
XX	MO200226997-A2.			
XX	04-APR-2002.			

XX 25-SEP-2001: 2001MO-US030124.
PF 26-SEP-2000: 2000US-0235487P.
XX (CEPT-) CEPTYR INC.
XX Luche RM, Wei B;
XX PI
XX WPI; 2002-315802/35.
DR N-PSDB; ABK48378.
XX
XX New DSP-16 polypeptide, useful for identifying modulators of its
PT activity, which can be used in the treatment of disorders such as
PT Duchenne muscular dystrophy, or cancer.
XX
XX Claim 50; Fig 4; 87pp; English.
XX
XX The present invention relates to a new polypeptide, DSP-16, having a 665
CC amino acid sequence, given in the specification, or a variant having at
CC least 50 % identical residues, which retains the ability to
CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
CC invention can be used for identifying agents which modulate DSP-16
CC activity, for modulation of a proliferative response in a cell, survival
CC of a cell, or differentiation of a cell. The cell displays contact
CC inhibition of cell growth or anchorage independent growth and may display
CC altered intercellular adhesion. The agent may modulate apoptosis, or the
CC cell cycle. The identified modulators can be used to treat Duchenne
CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
CC cell proliferation, and cell cycle abnormalities. The present amino acid
CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
CC alternative form protein of the invention. This sequence is encoded by
CC the human DSP-16 gene located on chromosome 12p
XX
XX SQ Sequence 517 AA;
XX
XX Alignment Scores:
Pred. No.: 0 Length: 517
Score: 472.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.76% Indels: 0
DB: 5 Gaps: 0
US-10-029-345a-108_COPY_532_2532 (1-2001) x AAU79159 (1-517)
QY 586 ACCGTCGCAAGCTGATCTTATCCCGAGTCTCATTTCTCGGTGCTGGAATGAC 645
DB 46 ThCyBProlyserProaspRheHleProGluSerHisPheLeuAlaGValProValaAsp 65
QY 646 AGCTTTGTGAGAAAATTTTGCSCGTGTTGAGCAAAATCAGTAGATTCTTGAAGAAGA 705
DB 66 SerPheCysGluIleuValIleuSerProTPrLeuAspIlySerValAspPheIleGluIlyVala 85
QY 706 AAAGCTCCCAATGAGATGTTCTTAGTGCACCTGTTAGCTGGGATCTCCGCTCCGACAC 765
DB 86 LysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThr 105
QY 766 ATGGCTATGCGCTACATCATGAGAGAGATGATGATGCTTTAGATGATGATGATGATGAT 825
DB 106 IleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 125
QY 826 GTGAAAGAAAAGACCTATATATCTCCAAATCTTAATTTTCTGGGCAACTCTTGAC 885
DB 126 ValIyGluIlyAsnArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuAsnArg 145
QY 886 TATGAGAGAAAGTTTAAAGAACGACTGAGCATCAGGGCCAAAGCAATCAAGCTG 945
DB 146 TyGluIlyLeuValIleLysAsnGlnThrGlyAlaSerGlyProLysSerIlyLeuIlyVal 165
QY 946 CTGCACCTGGAGAAAGCAATGAACCTGCTCCCTGCTCTCAGAGGATGGACAGAAAGC 1005

DB 166 LeuHisLeuGluIlyArgProAsnGluProValProAlaValaSerGluGlyGlnIlySer 185
QY 1006 GAGAGCCCCCTCAGTTCACCCCTGTGCGGACTCTGCTACTCAGAGGACAGCAAAAG 1065
DB 186 GluThrProLeuSerProCysAlaAspSerAlaThrSerGluAlaIleGlyGlnArg 205
QY 1066 CCCGTGCATCCCGCCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAG 1125
DB 206 ProValHisProAlaSerValProSerValProSerValGlnProSerLeuLeuGluAsp 225
QY 1126 ACCCGCTGTGATCAGAGCGCTCAGTGGCTGACCTGTCCGACAGACAGCTGGAAGACAG 1185
DB 226 SerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSer 245
QY 1186 AATTAAGCTTAACGCTTCTCTCTCTGATATCAATCAGTTTCATATTACGACAGCATG 1245
DB 246 AsnIlyLeuIlyArgSerPheSerLeuAspIleIlySerValSerIlySerAlaSerMet 265
QY 1246 GGAGCATCCTTACATGGCTTCTCCATGAGAAAGTCTTGGAACTACAAACCTTCC 1305
DB 266 AlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGluIlyTyrylyProSer 285
QY 1306 ACTACTCTGATGGAGCAACAAGCTATGCGAGTTCTCCCTGTTCCAGGAATATCGAG 1365
DB 286 ThrThrLeuAspGlyThrAsnIlyLeuCysGlnPheSerProValGlnGluLeuSerGlu 305
QY 1366 CAGACTCCCGAAACAGTCTCTATTAAGAGAAAGCCAGATCCCAAGAACTGCGACAC 1425
DB 306 GlnThrProGluThrSerProAspIlyGluGlnAlaSerIleProIlyLeuGlnThr 325
QY 1426 GCGAGCGCTTCAGACAGCGGAGCAAGGATTCGATTCGATTCGATTCGATTCGATTCGAT 1485
DB 326 AlaAspProSerAspSerIleSerIlyValGlyLeuHisSerValArgThrSerSerGly 345
QY 1486 ACCGCGCAAGAGTCCCTTTATCTCCACTGATCGATCGAAGTGGAGCGCTGAGGACAATTAC 1545
DB 346 ThrAlaGlnArgSerLeuLeuSerProLeuHisAspSerGlySerValGluAspAsnIly 365
QY 1546 CACACCAAGCTTCTTTTGGCTTTTCCACACCGACAGACAGCTTCAGAAAGTCTGCTGAC 1605
DB 366 HisThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeuThrIlySerAlaGly 385
QY 1606 CTGGGCTTAAAGGCTGCGACCTCGATATCTTGGCCCGCCAGACCTTACCCCTTCCCG 1665
DB 386 LeuGlyLeuIlyGlyIlyPheIleSerAspIleLeuAlaProGlnThrSerThrProSerLeu 405
QY 1666 ACCAGAGCTGATATTTTCCACAGAGTCTCTCACATTTACTCTGCTCAGCATCTAC 1725
DB 406 ThrSerSerIlyPheAlaThrGluSerSerHisPheIlySerHisAlaIleIly 425
QY 1726 GAGAGCAGTGCAGTTACTCTGCTTACAGCTGCAAGCAGCTGCCCATTTGCGAGACAA 1785
DB 426 GlyIlySerAlaSerIlySerIlyAlaTySerCysSerGlnLeuProThrCysGlyAspGln 445
QY 1786 GTCTATTCTGTCGCGAGCGGAGGAGAAAGCAATGACAGACGTCAGCTCGGGGAGACTGG 1845
DB 446 ValIlySerValArgArgIlyGlnIlyProSerAspArgAlaAspSerIlyArgSerIly 465
QY 1846 CATGAAGAGAGCCCTTTTAAAGCAGTTTAAAGCAGAAAGCTGCGCAATGAAATTTGGA 1905
DB 466 HisGluGluSerProPheGluIlyGlnPheIlyValArgArgSerCysGlnMetGluPheGly 485
QY 1906 GAGAGCATCATGACAGAAACGTCACGGAAGAGCTGGGAGAAAGTGGGACAGTCACTT 1965
DB 486 GluSerIleMetSerGluAsnArgSerArgGluGluLeuGlyLeuValGlyIlySerGlnSer 505
QY 1966 AGCTTTTGGGAGAGATGAAATCATTAGAGTCTCC 2001
DB 506 SerPheSerGlySerMetGluIleIleGluValSer 517
RESULT 4
AAE04834
ID AAE04834 standard; protein; 665 AA.

XX AAE04834;
AC
XX 10-SEP-2001 (first entry)
DT
XX
XX Human SGP002 phosphatase polypeptide.
XX
XX Human; SGP002 phosphatase polypeptide; phosphatase-related disease;
KW immune-related disorder; organ transplant rejection;
KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW metabolic disorder; hemetopietic cancer; mood disorder; cardiac;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW attention disorder; cognition disorder; psychotic disorder; cyostatic;
KW neurological disorder; virucide; nootropic; cerebroprotective; therapy;
KW neuroprotective; antibacterial; vulnerrary; tranquilizer; antiasthmatic;
KW hypotensive; immunosuppressive; antipruritic; analgesic; hypertensive;
KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; migraine; chromosome 12p11.1-p12.1.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Domain 1..173
FT /label=Catalytic_domain
FT 158..297
FT Domain /label=Phosphatase_domain
XX
XX WO200146394-A2.
XX
XX 28-JUN-2001.
XX
XX 21-DEC-2000; 2000WO-US034736.
XX
XX 21-DEC-1999; 99US-017325SP.
XX 28-DEC-1999; 99US-0175766P.
XX 25-JAN-2000; 2000US-0178078P.
XX 31-JAN-2000; 2000US-0179301P.
XX
XX (SUGEN) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;
PI Hill RJ, Flanagan P;
XX WPI; 2001-418058/44.
XX DR N-PSDB; AAD09492.
XX
XX Novel phosphatase polypeptide useful for treating cancers, immune-related
FT diseases and disorders, cardiovascular disease, brain or neuronal-
FT associated diseases and metabolic disorders.
XX
XX Claim 7; Fig 2; 186pp; English.
XX
XX The present invention relates to phosphatase polypeptides, nucleotide
CC sequences encoding them, as well as various products and methods useful
CC for the diagnosis and treatment of various phosphatase-related diseases
CC and conditions. Substance that modulates the activity of phosphatase
CC polypeptide is used to treat immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases and
CC metabolic disorders, including cancers of tissues, cancers of
CC hematopoietic origin, diseases of central and peripheral nervous system,
CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
CC lateral sclerosis, viral infections, infections caused by prions,
CC bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual
CC dysfunction, mood disorders, attention disorders, cognition disorders,
CC hypertension, hyperextension, psychotic disorders, neurological disorders,
CC dyskinesias and organ transplant rejection. The present amino acid
CC sequence is human SGP002 phosphatase polypeptide. This sequence is
CC classified as dual specificity phosphatase (DSP) and MAP kinase
CC phosphatase (MKP). SGP002 gene maps to chromosomal position 12p11.1-p12.1
XX
SQ Sequence 665 AA;

Alignment Scores:

Pred. No.: 0 Length: 665
Score: 472.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.76% Indels: 0
DB: 4 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x AAE04834 (1-665)

QY 586 ACTGTCCTCAAGCCTGACTTTATCCCGAGTCTCATTTCTGCGTGTGCTGTAATGAC 645
DB 194 TThCysProLysPheApsPheHleProGluSerHisPheLeuArgValProValAsnAsp 213
QY 646 AGCTTTTGAGAAAAATTTTGGCGGTTTGACCAAAATCAGTAGATTTCATTGAGAAAGCA 705
DB 214 SerPheCysGluLysIleLeuProTyrLeuAspLysSerValAspPheIleGluLysAla 233
QY 706 AAAGCCTCCAAATGATGTTGTTCTAGTGACATGTTAGCATGTTGCGGATCTCCGCTCCGAC 765
DB 234 LysAlaSerAenGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThr 253
QY 766 ATCGCTATCGCTCATCATGATGAAGAGATGACATGTTCTTATGATGAAGCTTACAGATT 825
DB 254 IleaIaIleaIa 273
QY 826 GTGAAGAAAAAAGACCTACTATATCTCAAACTTCAATTTCTGGGCGCAACTCCCTGAC 885
DB 274 ValLysGluLysArgProThrIleSerProAsnPheAsnLeuGluLeuLeuAsp 293
QY 886 TATGAGAAAGAAAGATTAAAGAACAGACTGAGCATCAGAGCCAAAGCAAACTCAAGCTG 945
DB 294 TyrGluLysLysIleLysAsenGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeu 313
QY 946 CTGACCTGAGAGAAAGCAAAATGAACCTGCTCCGTCTGTCAGAGGCTGACAGAAAGAC 1005
DB 314 LeuHisLeuGluLysProAsnGluProValProAlaValaSerGluIleGlyGlnLysSer 333
QY 1006 GAGACGCCCTCAGTCCACCTCTGTCGCACTGTCTACTACCTCAGAGGACAGACAAAG 1065
DB 334 GluThrProLeuSerProProCysAlaAspSerAlaThrSerGluAlaIleGlnThr 353
QY 1066 CCGGTGATCCCGCCAGCGCGTCCAGCGCCAGCGTCCAGCGTCTGTTAAGAGAC 1125
DB 354 ProValHisProAlaSerValProSerValProSerValGlnProSerLeuGluAsp 373
QY 1126 AGCCCGCTGTACAGAGCGCTCAGTGGGCTGACCTGTCCGACAGACAGGCTGGAAGACG 1185
DB 374 SerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSer 393
QY 1186 AATAGCTCAAGGCTTCTTCTCTCTGATATCAAAATCATGTTTCATATTACGACGACATG 1245
DB 394 AsnLysLeuLysArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMet 413
QY 1246 GCAGCATCTTTCATGAGCTTCTCTCATCAGAAAGATGTTTGAATATATAAACCCTCC 1305
DB 414 AlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGluLysTyrLysProSer 433
QY 1306 ACTACTCTGATGGAGCAACAAGCTATGCCAGTTTCTCCCTGTTCAGAACTATCGGAG 1365
DB 434 ThrThrLeuAspGlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGlu 453
QY 1366 CAGACTCCCGAAACAGAGTCTGATAGAGAGGAAGCCAGCATCCCAAGAAAGCTGACAGC 1425
DB 454 GlnThrProGluThrSerProAspLysGluAlaSerIleProLysLysLeuGlnThr 473
QY 1426 GCGAGGCTTGAACAGACGACAGCAAGCGATTGCGTCAAGAACACAGACAGAGTGGC 1485
DB 474 AlaArgProSerAspSerGlnSerLysArgLeuHisSerValAlaGlnThrSerSerGly 493
QY 1486 ACCGCCAGAGGTCCTTTTATCTTCACCTGCATGGAAGTGGAGCGTGAAGCAATTAC 1545
DB 494 ThrIaGlnArgSerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyr 513

QY 1546 CACACAGCTCTCTTTTCGSCCTTTCCACAGCCAGACGACCTCAAGATGCTGCGC 1605
 DB 514 HstHrSerPheLeuPheGlyLeuSerThSerGlnHstLeuThrYsSerAlaGly 533
 QY 1606 CTGGAGCTTAAGAGGCTGACACTCGATATCTTGAGCCCGCCAGACCTTACCTTCCTCG 1665
 DB 534 LeuGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 553
 QY 1666 ACCAGAGCTGTGATTTTGGCCACAGAGTCTCACTTCTGCTGCTGAGCATCTAC 1725
 DB 554 ThSerSerTrpYrPheAlaTrnGlnSerSerHstPheYrSerAlaSerAlaLeuYr 573
 QY 1726 GAGGAGAGGCGGCTTACTGCTGCTACAGTGCAGCCAGCTGCGCCCTTGGAGACCAA 1785
 DB 574 GlyGlySerAlaSerYrSerAlaYrSerCySerSerGlnLeuProThrCySGlyAaPgl 593
 QY 1786 GTCTATTTCTGTGGCAGGCGGCGAGAGCCAGAGTGAAGAGTGAAGTGAAGTGAAGTGA 1845
 DB 594 ValYrSerValArGArGArGlnYrPProSerAaPglAaPseArGArGArGArGArG 613
 QY 1846 CATGAAGAGAGCCCTTTGAAAAGAGTTTAAAGAGAGTGCAGAAATGAATTGGA 1905
 DB 614 HstGlnGlnSerProPheGlnYrGlnPheYrArGArGArGArGArGArGArGArG 633
 QY 1906 GAGAGCATCTATTCAGAAACAGGTCAAGGAAAGAGTGGGAAAGTGGGAGTCACTCT 1965
 DB 634 GlnSerIleuSerSerGlnAaArgSerArGlnGlnLeuLeuYrValGlySerGlnSer 653
 QY 1966 AGCTTTGGGCGAGCATGGAATCATTTAGAGTCTCC 2001
 DB 654 SerPheSerIleYrSerMetGlnIleGlnValSer 665
 RESULT 5
 AAU09016
 ID AAU09016 standard; protein; 665 AA.
 XX
 AC AAU09016;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human dual specificity phosphatase 21117.
 XX
 KW Human; dual specificity phosphatase 21117; hepatotropic; cytosolic;
 KW hematopoietic disorder; autoimmune disorder; diabetes mellitus;
 KW rheumatoid arthritis; multiple sclerosis; Crohn's disease;
 KW liver disorder; erythroid associated disorder; hemolytic anaemia;
 KW cellular proliferative; differentiative disorder; leukemia;
 KW acute myeloid leukemia; carcinoma; sarcoma; metastatic cancer;
 KW immunogen.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Domain location/Qualifiers
 FT 11..131
 FT /label= Rhodanese-like domain
 FT 158..297
 FT /label= Catalytic domain
 FT /note= "Dual specificity phosphatase catalytic domain"
 FT 242..254
 FT /label= Tyrosine-specific_protein_phosphatase_active_site
 FT
 XX
 XX MO200173059-A2.
 XX
 XX 04-OCT-2001.
 XX
 XX 23-MAR-2001; 2001MO-US009477.
 XX
 XX 24-MAR-2000; 2000US-0191858P.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Meyers RA;

XX
 DR MPI: 2001-611635/70.
 DR N-PSDB; AAS14639.
 XX
 PT New human dual specificity polypeptides and nucleic acids for diagnosis
 PT of disease and treatment of e.g. liver disorders.
 XX
 PS Claim 9; Fig 1; 143pp; English.
 XX
 CC The invention relates to two novel human dual specificity phosphatases
 CC designated 21117 and 38692, the nucleic acids encoding them (including
 CC fragments, allelic variants, their complements or nucleic acids that
 CC hybridise to them) and antibodies raised against the proteins. The
 CC antibody is useful for detecting the presence of the polypeptide, and the
 CC nucleic acid fragments are useful for detecting the presence of the
 CC nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
 CC antisense sequences) are useful for modulating the activity or expression
 CC of the polypeptide/nucleic acid, useful for the treatment of e.g.
 CC hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
 CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
 CC listed in the specification) liver disorders, erythroid associated
 CC disorders (e.g. hemolytic anaemia) cellular proliferative or
 CC differentiative disorders, leukemias (e.g. acute myeloid leukaemia),
 CC carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and
 CC 38692 are also useful for modulating the proliferation, survival,
 CC migration or differentiation of a 38692 or 21117-expressing cell. The
 CC polypeptide and nucleic acids are useful for identifying modulating
 CC agents. The present sequence represents the dual specificity phosphatase
 CC 21117
 XX
 SQ Sequence 665 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 665
 Score: 472.00 Matches: 472
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 70.76% Indels: 0
 DB: 4 Gaps: 0
 US-10-029-345A-108_COPY_532_2532 (1-2001) x AAU09016 (1-665)
 QY 586 ACCTGTCGAAGAGCTGATTTATCCCGAGTGTCAATTCCTGCGTGGATGAC 645
 DB 194 ThrCyProLysProAaPheLeuProGlnSerHstPheLeuArGValProValAaNap 213
 QY 646 ACCTTTTGAGAAAATTTTGGCGTGTGACCAATCAGTATTTCAATTGGAAGA 705
 DB 214 SerPheCyGlnYrIleLeuProTrpLeuAaPheValAaPheIleGlnYrAa 233
 QY 706 AAAGCTTCGAATGATGTGTTAGTCACTGTTAGCTGGAGATCTCCGCTCCGAC 765
 DB 234 LysAlaSerAaGlnGlyCyValLeuValHstCyLeuAlaGlyIleSerArGSerAlaTrn 253
 QY 766 ATCGCTATGCTTACATCATGAAGAGATGACATGTCCTTTAGATGAACCTTACAGATT 825
 DB 254 IleAlaIleAlaTrnIleMetLysArMetSerLeuAaPglAlaTrnYrArGpHe 273
 QY 826 GTGAAGAAGAAAAGACCTTATCTATCTCAAACTTCAATTTCTGGGCAACTCTGAGC 885
 DB 274 ValYrSGlnYrArGProThrIleSerProAaPheAaPheLeuGlnLeuLeuAaP 293
 QY 886 TATGAGAGAGATTAAGAACAGACAGTGAAGATCAGGCGCAAGAGCAAACTCAAGCTG 945
 DB 294 TyrGlnYrLeuYrIleYsAaGlnTrnGlyAlaSerGlyProYrSerYrLeuYrLeu 113
 QY 946 CTGCACTGGAAGAGCAAAATGAACCTGCTGCTGCTTCAAGAGGTGACAGAAAGC 1005
 DB 314 LeuHstLeuGlnYrProAaGlnYrProAlaValSerGlnGlyIleGlnYrSer 333
 QY 1006 GAGAGCGCCCTGAGTCAACCTGTGCGGAGTGTGACTCTCAAGAGCAGACAGCAAGG 1065
 DB 334 GlnTrnProLeuSerProProCyAaAaPseArAlaThSerGlnAlaIleGlnAaG 353

QY 1066 CCGTGATCCCGCAGCGTGCCAGCGTGCCAGCGTGCGTGTAAAGAC 1125
DB 354 ProValHisProAlaSerValProSerValGlnProSerLeuLeuLys 373
QY 1126 AGCCCCGTGTACAGGCGCTCAGTGGCTGCACCTGTCCGACAGAGCGTGGAGACAGC 1185
DB 374 SerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuLysSer 393
QY 1186 AATAAGCTCAGAGGCTCTCTCTCGGATCAATCAATGATTTTCAATTCAGCCAGCATG 1245
DB 394 AsnLysLeuLysArgSerPheSerLeuAspLysSerValSerTyrSerAlaSerMet 413
QY 1246 GACAGCATCTTTCATGAGCTTCTCTCATCAGAGAGATGCTTGGAAATACACAACTTCC 1305
DB 414 AlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSer 433
QY 1306 ACTACTCTGGATGGAGACCAAGCAAGCTATGCCAGTTCTCCCTGTTACAGAACTATCGAG 1365
DB 434 ThrThrLeuAspGlyThrAsnLysLeuLysGlnPheSerProValGlnGlnLeuSerGlu 453
QY 1366 CAGACTCCCGAAACCAAGTCTGATTAAGAGAGAGAGACATCCCGAAGAGCTGCACACC 1425
DB 454 GlnThrProGlnThrSerProAspLysGlnGlnAlaSerTyrProLysLysLeuGlnThr 473
QY 1426 GCCAGGCTTTCAGACAGCCAGACAGAGGATTCGATTCGATCAGAACCCAGCAGCAGTGC 1485
DB 474 AlaArgProSerAspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGly 493
QY 1486 ACCGCCAGAGAGTCCCTTTATCTTCCACGTGCATGAGAGAGCGGAGAGCAATTCAC 1545
DB 494 ThrAlaGlnArgSerLeuLeuSerProLeuHisArgSerGlySerValGlnAspAsnTyr 513
QY 1546 CACACACAGCTTCTTTTCGCGCTTTTCACACGACGACGACCTCAGAACTCTGCTGGC 1605
DB 514 HisThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGly 533
QY 1606 CTGGGCGCTTAAGGGCTGGCACTCGGATATCTTGCGCCCGCAGACCTTACCCCTTCCCTG 1665
DB 534 LeuGlyLeuLysGlyTyrPheHisSerAspLysLeuAlaProGlnThrSerThrProSerLeu 553
QY 1666 ACCAGCAGCTGTGATTTTGGCACAGAGTCTTCACACTTCTACTCTGCTCAGGCACTTAC 1725
DB 554 ThrSerSerTyrPyrPheAlaThrGlnSerSerHisPheLysSerAlaSerAlaIleTyr 573
QY 1726 GAGAGCAGTGCAGTTACTCTGCTTACAGCTGCAGCTGCGCCACTTGGAGAGACCA 1785
DB 574 GlyGlySerAlaSerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGln 593
QY 1786 GTCTATTCTGTGGCAGAGCGGACGAGACCAAGTGAAGAGCTGACTCGCGGCGAGCTGG 1845
DB 594 ValTyrSerValArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTyr 613
QY 1846 CATGAAGAGAGCCCTTTGAAAAGCAATTAAACGACAGAGCTGCAAAATGGAATTTTGA 1905
DB 614 HisGlnGlnSerProPheGlnLysGlnPheLysArgArgSerCysGlnMetClnuPheGly 633
QY 1906 GAGAGCATATGTCAGAGAAAGAGTCAACGGAAGAGCTGGGAAAGTGGCAGTCACTT 1965
DB 634 GlnSerIleLeuSerSerGlnAsnArgSerArgGlnGlnLysValGlyValGlySerGlnSer 653
QY 1966 AGCTTTTGGGAGCATGGAATCATTTAGAGTCTCC 2001
DB 654 SerPheSerGlySerMetGlnIleIleGlnValSer 665

RESULT 6

AAU79156
ID AAU79156 standard; protein; 665 AA.
XX
AC AAU79156;
XX
DT 02-JUL-2002 (first entry)
XX

DE Human dual-specificity phosphatase-3 (DSP-16) protein.
XX
KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW cancer; graft-versus-host disease; allergy; metabolic disease;
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
KW intercellular adhesion; DSP-16 modulator; chromosome 12p.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 242..251
FT /label= Active_site_domain
PN MO200226997-A2.
XX
PD 04-APR-2002.
XX
PP 25-SEP-2001; 2001WO-05030124.
XX
PR 26-SEP-2000; 2000US-0235487P.
XX
PA (CEPT-) CEPTYR INC.
XX
PI Luche RM, Wei B;
XX
DR WPI: 2002-315802/35.
DR N-PSDB; ABK47596.
XX
PT New DSP-16 polypeptide, useful for identifying modulators of its
PT activity, which can be used in the treatment of disorders such as
PT Duchenne muscular dystrophy, or cancer.
XX
PS Claim 1; Fig 2; 87pp; English.
XX
CC The present invention relates to a new polypeptide, DSP-16, having a 665
CC amino acid sequence, given in the specification, or a variant having at
CC least 50 % identical residues, which retains the ability to
CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
CC invention can be used for identifying agents which modulate DSP-16
CC activity, for modulation of a proliferative response in a cell, survival
CC of a cell, or differentiation of a cell. The cell displays contact
CC inhibition of cell growth or anchorage independent growth and may display
CC altered intercellular adhesion. The agent may modulate apoptosis, or the
CC cell cycle. The identified modulators can be used to treat Duchenne
CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
CC cell proliferation, and cell cycle abnormalities. The present amino acid
CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
CC protein of the invention. This sequence is encoded by the human DSP-16
CC gene located on chromosome 12p
XX
SQ Sequence 665 AA;
XX
Alignment Scores:
Pred. No.: 0 Length: 665
Score: 472.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.76% Indels: 0
DB: 5 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x AAU79156 (1-665)
QY 586 ACCTGTCGCAAGCGCTGATTTTCCCGAGTCTATCTTCTGCGTGCGCTGTGAATGAC 645
DB 194 ThrCysProLysProAspPheIleProGlnSerHisPheLeuArgValProValAsnAsp 213
QY 646 AGCTTTTGGAGAAATTTTGGCGGTGGACAAATAGTAGATTTTCAATTCAGAAAGCA 705
DB 214 SerPheCysGlnLysIleLeuProTyrPheAspLysSerValAspPheIleGlnLysAla 233

QY 706 AAAGCCTCAATGATGTGTTAGTCACTGTTTAACTGGAGATCTCCGCTCCGACAC 765
 DB 234 LyhAlaserbnglyCyvalleValHiCybeuHlaGlyIleSerArgSerAlaThr 253
 QY 766 ATGCTATGCGCTTACATCATGAAGAGAGAGACATGCTTTAGATGAAGCTTACAGATT 825
 DB 254 HleAlaIleAlaTyrIleMetIlyArgMetAspMetSerLeuSerbGluIaTyrArgPhe 273
 QY 826 GTGAAGAAAAAAGAACTACTATATCTCCAACTCAATTTTCTGGGCGCAACTCTTGAC 885
 DB 274 ValIySeGlyIySerArgProThrIleSerProAsnPhenLeuGlyGlnLeuLeuAsp 293
 QY 886 TATGAGAAGAAGATTAGAACAGACTGAGACATCAGAGCGCAAGAGCAAACTCAAGCTG 945
 DB 294 TyrGlyIyIySerIleIySerbGlnThrGlyAlaSerGlyProIySerIyIySerIy 313
 QY 946 CTGCACTGAGAGAGCCAAATGAACCTGTCCCTGCTCTCAGAGGGGTGACAGAAAAGC 1005
 DB 314 LeuHleIySeGlyIySerProAsnGlnProValProAlaValSerGlyGlnIyIySer 333
 QY 1006 GAGAGCGCGCTCACTCACTGCGGACTCTGCTTACCTCAGAGGAGAGAGCAAAAG 1065
 DB 334 GluThrProIySerProProCysAlaAspSerAlaThrSerGlyAlaIaGlyGlnArg 353
 QY 1066 CCGGTGATCCGCGCAGCGTGGCCAGCGTGGCCAGCGTGGCTGTTAGAGGAC 1125
 DB 354 ProValHisProAlaSerValProSerValProSerValGlnProSerLeuLeuGlnAsp 373
 QY 1126 AGCGCGCTGTACAGCGCTCAGTGGGCTCAGCTGCGCAGACAGCTGAGAGACAGC 1185
 DB 374 SerProIyValGlnAlaLeuSerGlyLeuHleIySerHlaAspArgLeuGlnIyAspSer 393
 QY 1186 AATTAAGCTCAAGCTTCTCTCTCTGATATCAATCACTGTTTCAATTCAGCAGATG 1245
 DB 394 AsnIyIySerIySerArgSerPheSerLeuAspIleIySerValSerIySerAlaSerMet 413
 QY 1246 GCAGCATCTTACATGCTTCTCTCTCATCAGAGAGATCTTGGAAATCTCAAACTCTCC 1305
 DB 414 AlaIaIaSerLeuHleGlyPheSerSerSerGlnAspAlaLeuGlnIyTyrIyIyProSer 433
 QY 1306 ACTACTGTGATGGAGCAACAAGCTATGCACTTCTCCCTGTTCAGAACTATCGAG 1365
 DB 434 ThrThrIySerbGlyThrAsnIyIyLeuCyGlnPheSerProValGlnIyIySerGln 453
 QY 1366 CAGACTCCGAAACCACTCTGATAGAGAGAGAGCCGACTCCCAAGAAAGCTGCGACAC 1425
 DB 454 GlnThrProGlnThrSerProAspIyGlnIyIyIaSerIleProIyIyIyLeuGlnThr 473
 QY 1426 GCCAGGCTTCAAGCAGCAGAGAGAGAGATGCTTGGTCAAGAACAGCAGAGTGGC 1485
 DB 474 AlaArgProSerArgSerGlnSerIyIyArgLeuHleIySerValArgIyIySerSerGly 493
 QY 1486 ACCGCCAGAGGTCCCTTTATCTCACTCATCGAAGTGGAGCGTGGAGAGCAATTAC 1545
 DB 494 ThrAlaGlnArgSerIyLeuSerProLeuHleArgSerGlySerValGlnIyAspArgTyr 513
 QY 1546 CACACCAAGCTTCTTTTGGGCTTTCCACAGCCAGAGAGACACTCAGAAAGTGTGTCGC 1605
 DB 514 HisThrSerPheLeuPheGlyIyLeuSerThrSerGlnGlnHisLeuThrIyIySerAlaGly 533
 QY 1606 CTGGGCGCTTAAGGCTGGACCTGGGATATCTGGGCGCCAGAGCTTACCCCTGCTCC 1665
 DB 534 LeuGlyIyIySerGlyIyPheIyIySerAspIleLeuIaIaProGlnThrSerIyIyProSerLeu 553
 QY 1666 ACCAGCAGCTGATTTTGGCAGAGAGTCTTCACTTACTTCTGCTCAGCCATCTAC 1725
 DB 554 ThrSerSerIyIyPheAlaIaThrGlnSerSerHisPheIyIySerAlaIaIaIeTyr 573
 QY 1726 GAGAGCGTGGCACTTCTGCTTACAGCTGACGAGCAGCTGCCCACTTGGGAGAGACAA 1785
 DB 574 GlyIyIySerAlaIaSerIyIySerAlaIaIySerCysSerGlnLeuProThrCysGlyIyAspGln 593
 QY 1786 GTCTATTCTGTGGCGAGCGGAGAGCAAGCTAGTACAGACTGACTGCGCGGAGCTGG 1845

DB 594 ValTyrSerValArgArgArgGlnIyIyProSerAspArgAlaAspSerArgTyr 613
 QY 1846 CATGAAGAGAGCCCTTTGAAAAAGCATTAAACGAGAAAGCTCCCAATGAAATTGGA 1905
 DB 614 HisGlnGlnSerProPheGlnIyIySerGlnPheIyIyArgArgSerCysGlnMetGlnPheGly 633
 QY 1906 GAGAGCATATGTCAGAGAAAGGTCACGAGAAAGACTGGGCAAGAGTGGGCACTGCTC 1965
 DB 634 GluSerIleIySerGlnIyIySerArgGlnIyIySerGlnIyIyValGlySerGlnSer 653
 QY 1966 AGCTTTCCGCGCAGCATGGAATCATTTAGGCTCC 2001
 DB 654 SerPheSerGlySerMetGluIleIleGluValSer 665
 RESULT 7
 AAU09946 standard; protein; 665 AA.
 AAU09946:
 18-UTN-2002 (first entry)
 DE Protein sequence of human (dual specificity phosphatase) DUSP-10.
 KW Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke;
 KW neuronal degeneration syndrome; Alzheimer's disease; depression;
 KW schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
 KW osteoporosis; diabetes.
 OS Homo sapiens.
 PN WO200177340-A1.
 PD 18-OCT-2001.
 PF 06-APR-2001; 2001WO-BP003966.
 PR 10-APR-2000; 2000EP-00107143.
 PA (MERE) MERCK PATENT GMBH.
 PI Duecker K;
 DR WPI, 2002-010917/01.
 DR N-PSDB; AAS15768.
 XX Novel dual specificity phosphatase polypeptides useful for treating
 PT cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
 PT disease, depression, schizophrenia, asthma and immune disorders.
 XX
 PS Claim 2; Page 37-39; 43pp; English.
 CC The present invention relates to a new isolated dual specificity
 CC phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
 CC sequence that is fully defined in the specification. The invention also
 CC provides a sequence encoded by a 3059 nucleotide sequence fully defined
 CC in the specification, and a sequence having at least 95 % identity to the
 CC polypeptide, or fragments or variants of DUSP-10. The invention is useful
 CC for treating cancer e.g. leukemia, colon carcinoma, lung cancer,
 CC prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
 CC stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
 CC schizophrenia, cardiac myopathies, asthma, immune disorders,
 CC inflammatory processes e.g. arthritis, bowel disease, type I diabetes,
 CC osteoporosis, diabetes and diabetes associated diseases. The molecules of
 CC the invention are also useful as vaccines for inducing immunological
 CC response in a mammal, in disease diagnosis and in assays for screening
 CC agonistic or antagonistic compounds. Other uses of the invention include
 CC identifying membrane bound or soluble receptors, as a diagnostic reagent,
 CC in chromosome localisation studies, and as a valuable tool in tissue
 CC expression studies. The present sequence represents the dual specificity
 CC phosphatase, DUSP-10, protein of the invention
 XX

XX 07-FEB-2002.
 PD
 XX
 PF 26-JUL-2001; 2001MO-US023716.
 XX
 PR 28-JUL-2000; 2000US-0221679P.
 PR 03-AUG-2000; 2000US-0223372P.
 PR 10-AUG-2000; 2000US-0224309P.
 PR 18-AUG-2000; 2000US-0226728P.
 PR 30-AUG-2000; 2000US-0229254P.
 PR 08-SEP-2000; 2000US-0231366P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 PI Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;
 PI Stewart BA, Gandhi AR, Patterson C, Lee EA, Hatalia AD, Lu DM,
 PI Tribouley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;
 PI Walla NK, Kearney L;
 XX
 DR MPI: 2002-188735/24.
 DR N-PSDB; ABK14474.
 XX
 PS
 XX
 PT New protein phosphatases, useful for diagnosing, treating or preventing
 PT immune system disorders (e.g. Crohn's disease), neurological disorders
 PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
 PT cancers).
 XX
 PS Claim 1; Page 107-108; 117pp; English.
 XX
 CC The present invention relates to a new polypeptide, a naturally occurring
 CC amino acid sequence at least 95 % identical to it, a biologically active
 CC fragment of it or an immunogenic fragment of it. The polypeptides,
 CC polynucleotides, agonists and antagonists are useful for diagnosing,
 CC treating or preventing disorders associated with aberrant expression of
 CC protein phosphatases (PP), particularly immune system disorders e.g.
 CC acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
 CC asthma or Crohn's disease, neurological disorders e.g. epilepsy,
 CC Huntington's disease, dementia or Parkinson's disease, developmental
 CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
 CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
 CC or sarcoma. The present amino acid sequence represents human protein
 CC phosphatase 7 (PP7) which is one of several human protein phosphatases
 CC (AAU5783-AAU5792) of the invention
 XX
 XX
 SQ Sequence 665 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 665
 Score: 472.00 Matches: 472
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 70.76% Indels: 0
 DB: 5 Gaps: 0
 US-10-029-345a-108_COPY_532_2532 (1-2001) x AAU5789 (1-665)
 QY 586 ACCTGTCGAAAGCTGATTCATCCCGAGCTTCATTCCTGCGTGTGCTGGAATGAC 645
 DB 194 ThCgProlsProlsPheleleProgluSerHisPheleuArgValProValaAsp 213
 QY 646 AGCTTTTGTGAGAAATTTTGGCGGTGTGGACAATCATAGTATTTCATTGAGAAAGCA 705
 DB 214 SerPheCysGluSerHisPheleuProTyrPheleuAspHisPheleleGluVala 233
 QY 706 AAGGCTTCAGATGATGTTCTAGTACCTGTTTGAAGTCCGCTCCGCGACAC 765
 DB 234 LysHisSerHisGluSerValHisPheleuValHisPheleuValHisPheleuValHis 253
 QY 766 ATGCTATGCGCTCATCATGAGAGAGATGACATGCTCTTTAATGAGAGCTTACAGATT 825
 DB 254 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 885
 QY 826 GTGAAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 885

DB 274 ValLysGluLysArgProThrHisSerProAlaPheAsnHisPheLeuGluGlnLeuLeuAsp 293
 QY 886 TATGAGAAAGAAATTTAAGAACCAAGATGAGATGAGGCGCAAGCAAACTCAAGCTG 945
 DB 294 TyrGluLysValHisPheleuSerHisPheleuSerHisPheleuSerHisPheleuSer 313
 QY 946 CTGCACTGAGAGAGCAAAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1005
 DB 314 LeuHisPheGluLysPheleuSerHisPheleuSerHisPheleuSerHisPheleuSer 333
 QY 1006 GAGAGCGCCCTGAGTCCAGCCCTGTCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065
 DB 334 GlnThrProLeuSerProProCysAlaPheSerHisPheleuSerHisPheleuSerHis 353
 QY 1066 CCCGTCATCCGCGCCAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1125
 DB 354 ProValHisPheleuSerHisPheleuSerHisPheleuSerHisPheleuSerHisPhe 373
 QY 1126 AGCGCGCTGATCAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCG 1185
 DB 374 SerProLeuValGlnHisPheleuSerHisPheleuSerHisPheleuSerHisPheleu 393
 QY 1186 AATAAGCTCAAGCGCTTCTCTCTGATATCAATCAGTTTCATATTCAGCCAGCATG 1245
 DB 394 AsnLysPheLysArgSerHisPheleuSerHisPheleuSerHisPheleuSerHisPhe 413
 QY 1246 GCAGCATCTTACATGAGCTTCTCTCATGAGAGATGCTTGGATATCTACAACTTCC 1305
 DB 414 AlaAlaSerLeuHisGluPheSerSerSerGluAspAlaLeuGluLysTyrLysProSer 433
 QY 1306 ACTACTCTGAGAGGAGCAAGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1365
 DB 434 ThrThrLeuAspGluLysHisPheleuSerHisPheleuSerHisPheleuSerHisPhe 453
 QY 1366 CAGACTCCCGAAGCAAGCTCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1425
 DB 454 GlnThrProGlnThrSerProAspLysGluGluAlaSerHisPheleuSerHisPheleu 473
 QY 1426 GCGAGGCTTCAAG 1485
 DB 474 AlaArgProSerAspSerGlnSerLysValHisPheleuSerHisPheleuSerHisPhe 493
 QY 1486 ACCGCGCAAGGCTCTTATCTCCAGTCGATCGAAGTGGAGCGTGGAGCAATTTC 1545
 DB 494 ThrAlaGlnArgSerLeuSerProLeuHisPheSerGluSerValGluAspAsnTyr 513
 QY 1546 CACACAGCTCTCTTTCGCGCTTTCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1605
 DB 514 HisThrSerPheleuPheGluLeuSerHisPheleuSerHisPheleuSerHisPheleu 533
 QY 1606 CTGGGCTTAAAGGCTGAGCTCGATATCTTGGCCCGAGAGAGAGAGAGAGAGAGAGAG 1665
 DB 534 LeuGluLysLeuGluLysHisPheleuSerHisPheleuSerHisPheleuSerHisPhe 553
 QY 1666 ACCAGAGCTGATTTTTCGAG 1725
 DB 554 ThrSerSerTrpTyrPheHisPheleuSerHisPheleuSerHisPheleuSerHisPhe 573
 QY 1726 GAGAGCAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1785
 DB 574 GlyLysSerHisPheleuSerHisPheleuSerHisPheleuSerHisPheleuSerHis 593
 QY 1786 GTCTATTCGTGCGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1845
 DB 594 ValTyrSerValArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrp 613
 QY 1846 CATGAAG 1905
 DB 614 HisGluLysSerProPheGluLysGlnPheLysValArgArgSerGlnPheGluPheGlu 633
 QY 1906 GAGAGCATGTCAG 1965

Dd		594	ValTyrSerValArgArgArgGlnIlyrProSerAspArgAlaAspSerArgArgSerTrp	613
Oy		1846	CATGAAGACGCCCTTTGGAAAGCATTTTAACGCAGAAGCTGCCAATGCAATTGTGA	1905
Dd		614	HtGtGtGtSerProPhgHlWlyeGInPheLysArgArgSerCyGInMetGlUpHeGly	633
Oy		1906	GAGAGCATTTCATCTCGAACAACGGTCACGGGAAGAGCTGGGGGAAGTGCGCACATCT	1965
Dd		634	GlSerIleMeSerGluAsnArgSerArgGtGtGtLeuGlyLysValGlySerIlnSer	653
Oy		1966	AGCTTTCCGGCAGCATGAAATCATTTAGAGTCTCC	2001
Dd		654	SerPheSerGlySerMetGluIleIleGluValSer	665
RESULT 10				
ID	AAU79929		standard; protein; 665 AA.	
XX	AAU79929;			
XX	02-JUN-2002	(first entry)		
DE			Human dual specificity phosphatase 21117 protein.	
KX			Human; dual specificity phosphatase 21117; erythroid-related disorder;	
KM			haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;	
KW			erythrocytosis; liver-related disorder; cancer.	
XX			Homo sapiens.	
OS				
XX				
FH	Key	Location/Qualifiers		
FT	Domain	11..131		
FT		/label= Rhodanese_1like_domain		
FT	Region	21..24		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Region	91..94		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Domain	158..297		
FT		/note= "Dual specificity phosphatase catalytic domain"		
FT	Region	214..217		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Active-site	242..254		
FT		/note= "Tyrosine specific protein phosphatase active site"		
FT	Region	242..254		
FT		/note= "C-X5-R motif"		
FT	Region	266..269		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Region	369..372		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Region	421..424		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Region	434..437		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Region	458..461		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Region	508..511		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Region	589..592		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Region	612..615		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Region	617..620		
FT		/note= "Casein kinase II phosphorylation site"		
FT	Region	642..645		
FT		/note= "Casein kinase II phosphorylation site"		
XX				
PN	US2002034807-A1.			
XX				
PD	21-MAR-2002.			
PF	23-MAR-2001; 2001US-00816494.			
XX				

PR	24-MAR-2000; 2000US-0191858P.
XX	(MEYE/) MEYERS R A.
PA	
XX	Meyers RA;
PI	
XX	WPI: 2002-351086/38.
DR	N-PsDB; ABK49402.
XX	New nucleic acids, designated 38692 and 21117, encoding dual specificity phosphatases for treating cell proliferation and differentiation disorders including hematopoietic and erythroid-related disorders and cancers.
PT	
XX	Claim 8; Fig 1; 76pp; English.
PS	
XX	The present invention relates to new nucleic acids designated 38692 and 21117 encoding dual specificity phosphatase family members. The nucleic acid, polypeptide encoded by it, and antibody specific for the polypeptide may be used to diagnose and treat haematopoietic-related disorders such as leukaemias and autoimmune diseases, erythroid-related disorders such as anaemias and erythrocytosis, liver-related disorders, and cancers, particularly of the breast, colon, adipose, prostate and lung. The present amino acid sequence represents the human dual specificity phosphatase 21117 protein of the invention, as described above
CC	
CC	
SQ	Sequence 665 AA;
	Alignment Scores:
Pred. No.:	0 Length: 665
Score:	472.00 Matches: 472
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	70.76% Indels: 0
DB:	5 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x AAU79929 (1-665)	
QY	586 ACCGTGCCAAAGCCTGACTTATCCCGAGCTCATTTCCGTGCCTGTGATGAC 645
Db	194 ThrcyProlyspRobaphneilleProclnSerlsrPhelenuArgValProvalAsnasp 213
QY	646 AGCTTTTGTAAGAAAATTTGGCCGTGGTGGACAATACTAGTAGATTTCATTGAGAAGCA 705
Db	214 SerPheCygluylsrleLeuProTrlpheapryservalAspPheilleglulysAla 233
QY	706 AAAGCTTCATGATGATGTGTCTAGTGACAGCTGTTTACTGGAGATCTCCCGTCCGCCACC 765
Db	234 LyslAserlAsengllyCyValleuValHlsybleuAlaglylserArsgserAlaThr 253
QY	766 ATCGGTATCGGCTCATCATGATAAGAGGTGACATGCTTTAGATGAAGCTTACAGATT 825
Db	254 lleaAlaileAlayrlleueLysarlgwchaspheiserleuabepGlualAyraIsgPhe 273
QY	826 GTGAAGAAAAAAGACCTACTATATCTCCAACCTTCAAATTTCTGGGCCAACTCTGAC 885
Db	274 VallysgluylsVatgPrtotHlleSerProasnPhenPheuleuglInleuLeuasp 293
QY	886 TATGAAGAAAGATTAAAGAACCAAGCTGGAGCATCAGGGCCAAAGACCAACTGACCTG 945
Db	294 TyrgluylslsrlylelYsaenglnThcglYAlaserglProlysserlySleulylslu 313
QY	946 CTGCAACCTGGAGGAAGCCAAATGAACCTGTCCCTGCTGTCTCAGAGGGGTGACAGAAAGC 1005
Db	314 LeuhlsleuglulysPProahenglpProvalProAlalaserglulysglInlulysSer 333
QY	1006 GAGAGCGCCCTCAGTCCACCTGTGCCGACTCTGCTCACTCAGAGCGACGACCAAGG 1065
Db	334 GlulnrPrroleusenSerProprocYAlaAspservalArhsersgluaAlaglYglndarg 353
QY	1066 CCGGTGATCCCGCCAGCGTGGCCAGGCTGCCAGCGTGCAGCGCTGCTTTAGAGGAC 1125

Db 354 ProValHisProAlaSerValProSerValProSerValGlnProSerLeuLeuGlnAasp 373
QY 1126 AGCCCGCTGGTACAGGCGCTCACTGGGCTGCACCTGTCCGACAGACAGGCTGGAAAGACAGC 1185
Db 374 SerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGlnAaspSer 393
QY 1186 AATAAGCTCAAGGCTCTCTCTGATATCAATTCAGTTTCATATTCAGGCACAGC 1245
Db 394 AsnLysLeuLysArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMet 413
QY 1246 GCAGCATCCTTACATGGCTTCTCCTCATCAGAAATGCTTGGAAATATACAAACCTTCC 1305
Db 414 AlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGlnTyrTyrLysProSer 433
QY 1306 ACTACTCTGGATGGACCAACAAAGCTATGCCAGTTCTCCCTGTTCAAGAACTATGGAG 1365
Db 434 ThrThrLeuAspGlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGln 453
QY 1366 CAGACTCCCGAAACAGTCTCTGATTAAGGAGGAGACGACATCCCGCAAGAACTGGACAC 1425
Db 454 GlnThrProGlnThrSerProAspLysGlnGlnAlaSerIleProLysLysLeuGlnThr 473
QY 1426 GCCAGCGCTTCAGACAGCCAGACCAAGCATTCGATTCGGTCAAGAACAGACAGCTGGC 1485
Db 474 AlaArgProSerAspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGly 493
QY 1486 ACCGCCCAAGAGTCCCTTTATCTCCATCTGATCGAAAGTGGAGCGCTGGAGCAATTAC 1545
Db 494 ThrAlaGlnArgSerLeuSerProLeuHisArgSerLysSerValGlnAaspArgTyr 513
QY 1546 CACACAGACTCTCTTTTCGGCTTTCCACACGACGACAGACCTCCACGAACTGCTGGCAG 1605
Db 514 HisThrSerPheLeuHeuHeuGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGly 533
QY 1606 CTGGGCGTTAAGGCGCTGGCACTGGAATCTTGCCCCCGACACCTTACCCCTTCCCTG 1665
Db 534 LeuGlyLeuLysGlyTyrPheHisSerAspIleLeuAlaProGlnThrSerThrProSerLeu 553
QY 1666 ACCAGAGCTGGATTTTGGCCACAGAGTCTTCACACTTTCATCTGCTCAGCATCTAC 1725
Db 554 ThrSerSerTyrTyrPheAlaThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyr 573
QY 1726 GGAGCAGTGGCCAGTTACTCTGCTCAGCTGACGTGACGAGCTGGCCACTGGCGAGACCA 1785
Db 574 GlyGlySerAlaSerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAaspGln 593
QY 1786 GTCTATTCTGTGGCAGCGCGGACGAACCAAGTACAGAGCTGACTCGCGCGGAGCTGG 1845
Db 594 ValTyrSerValArgArgArgGlnLysProSerAspArgAlaAspSerArgAspSerTyr 613
QY 1846 CATGAAGAGAGCCCTTTGAAAAGCAGTTTAAACGAGAAAGCTGCCAATGGAATTGGGA 1905
Db 614 HisGlnGlnSerProHeuGlnLysGlnPheLysArgArgSerCysGlnMetGlnPheGly 633
QY 1906 GAGAGCATCATGTCAAGAAACAGGTCAACGGAAGAGACTGGGAGAACTGGGACGTGCT 1965
Db 634 GluSerIleMetSerLeuAsnArgSerArgGlnGlnLeuGlyLysValGlyLysGlnSer 653
QY 1966 AGCTTTTGGGAGCATGGAATCATTTAGAGTCTCC 2001
Db 654 SerPheSerGlySerMetGlnIleIleGlnValSer 665

RESULT 11
ABB97291
ID ABB97291 standard; protein; 665 AA.
XX
AC ABB97291;
XX
DT 27-JUN-2002 (first entry)
XX
DE Novel human protein SEQ ID NO: 559.
XX
KW Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;

KW antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag.
OS Homo sapiens.
XX
PN MO200222660-A2.
XX
PD 21-MAR-2002.
XX
PF 10-SEP-2001; 2001MO-US026015.
XX
PR 11-SEP-2000; 2000US-00659671.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F,
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-292408/33.
DR N-PSDB; ABN32477.
XX
XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.
XX
PS Example 2; SEQ ID NO 559; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention
XX
SQ Sequence 665 AA:
XX
Alignment Scores:
Pred. No.: 0 Length: 665
Score: 472.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.76% Indels: 0
DB: Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x ABB97291 (1-665)
QY 586 ACCTGTCCAAAGCCCTGACTTATCCCGAGTCTCATTTCCGCGGTGCTGAATGAC 645
Db 194 ThrCysProLysProAspPheIleProGlnSerHisPheLeuArgValProValAsnAasp 213
QY 646 ACCTTTTGGAGAAATTTTGGCGTGTGGCAAAATGATGATTTTATTTGAGAAACA 705
Db 214 SerPheCysGlnLysIleLeuProThrPheAspLysSerValAspPheIleLysAla 233
QY 706 AAAGCTCCCAATGATGTGTTCTAGTGACATGTTAGCTGGGATCTCCGCTCCGCCAC 765
Db 234 LysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThr 253
QY 766 ATCGCTATCGCTTAATCATGAAAGGATGACATGCTCTTATGATGAAGCTTACGATTT 825
Db 254 IleAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAspGlnAlaTyrArgPhe 273
QY 826 GTGAAGAAAAAAGACTACTATATCTCCAAACTCAATTTTCTGGGCCAACTCTGGAC 885
Db 274 ValLysGlnLysArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuAasp 293
QY 886 TATGAGAAAGATTTAAGAACCAAGCTGAGCATGAGGCGCAAGAGCAAACTCAAGCTG 945
Db 294 TyrGlnLysLysIleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeu 313

```

QY 946 CTGCACCTGGAGAGCAAAATGAACCTGCTCCCTGCTCTCAAGAGGGGTGACAGAAAAGC 1005
   |||||
Db 314 LeuH1SerLeuLysPProAengLProValPProValSerGluGlnLysSer 333
QY 1006 GAGACGCCCTCAGTCCACCTGTGCGACTGTGCTCACTCAGAGCGACGAGCAAAAGG 1065
   |||||
Db 334 GluThrProLeuSerProProCybAlaAspSerLathSerGluAlaGlnArg 353
QY 1066 CCCGTGATCCCGCCAGCCGTCCCAAGCTGCGCCAGCTGACGCCCTGCTTGAAGAGAC 1125
   |||||
Db 354 ProValH1PProAlaSerValPProSerValGlnProSerLeuLeuGluAsp 373
QY 1126 AGCCCGCTGGTACAGCGGCTCAGTGGGCTGACCTGCGCAGACAGCTGGAAGACAGC 1185
   |||||
Db 374 SerProLeuValGlnAlaLeuSerGlyLeuH1SerL1aAspArgLeuGluAspSer 393
QY 1186 AATAAGCTCAAGCGTCTCTCTCTGATATCAATCAGTTTCATATTCAGCCAGCATG 1245
   |||||
Db 394 AsnLysLeuLysArgSerPheSerLeuAspL1eLysSerValSerLysSerL1aSerMet 413
QY 1246 GCAGCATCTTACATGGCTTCTCTCTCATCAGAAAGATCTTTGGAATPCTACAACTTCC 1305
   |||||
Db 414 AlaAlaSerLeuH1SgLYPheSerSerSerGluAspAlaLeuGluLysTYrTYrLysProSer 433
QY 1306 ACTACTGCTGGATGGGACCAACAAAGCTATGCGCAGTTCCTCCCTGTTCAGAACTATCGGAG 1365
   |||||
Db 434 ThrThrLeuAspGlyThrAsnLysLeuCybGlnPheSerProValGlnGluLeuSerGln 453
QY 1366 CAGACTCCCGAAACCACTCTGATTAAGAGAGAACCCAGCATCCCAAGAAAGCTGACAGCC 1425
   |||||
Db 454 GlnThrProGluThrSerProAspLysGluGluAlaSerL1eProLysLysLeuGlnThr 473
QY 1426 GCCAGGCTTTCAGACAGCCAGAGCAGATTGCAATTGCGTCAAGAACCCAGCAGCATGTCG 1485
   |||||
Db 474 AlaArgProSerAspSerGlnSerLysArgLeuH1SerValArgThrSerSerSerGly 493
QY 1486 ACCGCCAGAGGTCCTTTATCTCCATCGATCGAAGTGGGCGCGTGGAGGACATTAAC 1545
   |||||
Db 494 ThrAlaGlnArgSerLeuLeuSerProLeuH1ArgSerL1eSerValGluAspAsnTYr 513
QY 1546 CACAGCAGCTTCTCTTTCGCTTTCACAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1605
   |||||
Db 514 HisThrSerPheLeuPheGlyLeuSerThrSerGlnH1SerL1eThrLysSerAlaGly 533
QY 1606 CTGGGCTTAAAGGCTGCGACTCGGATATTTGGCCCCCAGAACCTTACCTTCCCTG 1665
   |||||
Db 534 LeuGlyLeuLysGlyTrpHisSerAspL1eLeuAlaProGlnThrSerThrProSerLeu 553
QY 1666 ACCAGCAGCTGATATTTTGCACAGAGTCTCTCACTTCACTCTGCTGACCATCTAC 1725
   |||||
Db 554 ThrSerSerTrpLysPheAlaThrGlnSerSerHisPheLysSerAlaSerAlaL1eTYr 573
QY 1726 GAGAGCAGTCCAGATTACTGTGCTACAGCTGACGAGCAGCAGCTGCGCACTTCCGAGAGCA 1785
   |||||
Db 574 GlyGlySerL1aSerLysSerL1eTYrSerCybSerGlnLeuProThrCybGlyAspGln 593
QY 1786 GTCTATTCTGTGCGCAGCGGCGAGAACCAAGTACAGAGCTGACTCGCGCGAGAGCTGG 1845
   |||||
Db 594 ValLysSerValArgArgArgGlnLysProSerAspArgL1aAspSerArgLysSerTrp 613
QY 1846 CATGAAGAGAGCCCTTTGAAAAAGCAGTTTAAACGCGAAGAGCTGCCAAATGAAATTGCA 1905
   |||||
Db 614 HisGlnGlnSerProPheGlnLysGlnPheLysArgArgSerCybGlnMetGluPheGly 633
QY 1906 GAGAGCATCATGTCAAGAGACAGCTCAAGGAGAGAGCTGGGAGAAAGTGGCAGTCACTG 1965
   |||||
Db 634 GluSerL1eMetSerGluAsnArgSerArgGluGluLeuGlyValGlySerGlnSer 653
QY 1966 AGCTTTTGGGAGAGCATGAAATCATTTAGAGTCTCC 2001
   |||||
Db 654 SerPheSerGlySerMetGluL1eL1eGluValSer 665

```

```

RESULT 12
ADA54744
ID ADA54744 standard; protein; 665 AA.
XX
AC ADA54744;
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 2312.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW Inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-0006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
XX
PR 24-JAN-2002; 2002US-0350435P.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y;
XX
DR WPI, 2003-395539/38.
DR N-PDB; ADA53105.
XX
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory in
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14, SEQ ID NO 2312; 205pp; English.
XX
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54072). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 665 AA;

```

Alignment Scores:

Pred. No.:	Length:
0	665
Score: 472.00	Matches: 472
Percent Similarity: 100.00%	Conservative: 0
Best Local Similarity: 100.00%	Mismatches: 0
Query Match: 70.76%	Indels: 0
DB: 6	Gaps: 0

```

US-10-029-345A-108_COPY_532_2532 (1-2001) x ADA54744 (1-665)
QY 586 ACCGTGCCAAAGCTGATCTTATCCCGAGTGCATTTCTGCGTGGCTGGAATGAC 645
   |||||
Db 194 ThrCybProLysProAspPheL1eProGlnSerHisPheLeuArgValProValAsnAsp 213
QY 646 AGCTTTTGAGAAAATTTTGCCTGTTGACCAATCAGTATTTCAATTGAGAAAGCA 705
   |||||
Db 214 SerPheCybGluLysL1eLeuProTrpLeuAspLysSerValAspPheL1eGluLysAla 233
QY 706 AAAGCTCCAAATGATGTGTTCTAGTGCACCTGTTTACCTGGGATCTCCCGCTCCGACAC 765
   |||||
Db 234 LysAlaSerAsnGlyCybValLeuValHisCybLeuAlaGlyL1eSerArgSerAlaThr 253
QY 766 ATCGCTATGCGCTACATCATGGAAGAGGATGAGCATGCTTTTGAATGATTAACATTT 825
   |||||

```


Alignment Scores:

Pred. No.: 0
 Score: 452.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 67.77%
 DB: 5
 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x AAU79161 (1-665)

```

QY 646 AGCTTTTGAGAAATTTTCCGCTGTGACAAATCAGTAAATTCATTGAGAAAGCA 705
DB 214 SerPheCysGluYsrIleuProThrLeuAspIleuSerValAspPheIleGluYsrAla 233
QY 706 AAGGCTCCCAATGATGTGTTCTAGTCACTGTTTACCTGGATCTCCCGCTCCGACAC 765
DB 234 LysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerAlaGSerAlaThr 253
QY 766 ATGCTATCCGCTACATCATGAGAGATGAGATGCTTTAGATGAAGCTTACAGATT 825
DB 254 ILeAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 273
QY 826 GTGAAAGAAAAAGACCTACTATATCTCAAACTTCAATTTTCTGGCCCACTCTGAC 885
DB 274 ValIyGluYsrArgProThrIleSerProAsnPheAsnIleuGluYsrIleuLeuAsp 293
QY 886 TATGAGAAAGAAATTAAGAACCACTGAGCATCAGAGGCCAAAGCAAACTCAAGCTG 945
DB 294 TyrGluYsrIleYsrIleYsrAsnGluYsrIleYsrIleYsrIleYsrIleYsrIleYsr 313
QY 946 CTGACCTGTGAGAGAACCAATGAACCTGCTCTGCTGTCTCAGAGGGGTGACAGAAAGC 1005
DB 314 LeuHisLeuGluYsrProAsnGluYsrProAlaIleAlaSerGluYsrIleYsrIleYsr 333
QY 1006 GAGAGCGCCCTCAGTCACCTGTGCGACTGTCTCACTCAGAGGACGAGCAAAAG 1065
DB 334 GluThrProLeuSerProProCysAlaAspSerAlaThrSerGluYsrIleYsrIleYsr 353
QY 1066 CCCGTGATCCCGCCAGCGTCCCGCCAGCGTCCAGCGTCCAGCGTCCGTTTGAAGAC 1125
DB 354 ProValHisProAlaSerValProSerValProSerValGlnProSerLeuGluYsr 373
QY 1126 AGCCGCTGTATCAGAGGCTCAGAGGCTCAGCTGCGCAGACAGCGCTGGAAGACAG 1185
DB 374 SerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluYsr 393
QY 1186 AATAAGCTCAAGGCTTCTCTCTGTGATATCAATCAGTTTCATATTCAGCCAGCATG 1245
DB 394 AsnIleuLeuYsrArgSerPheSerLeuAspIleYsrSerValSerIleYsrAlaSerMet 413
QY 1246 GCGCATCTTATCATGAGCTTCTCTCATCAAGAGTCTTTGAATTCATCAAACTTCC 1305
DB 414 AlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGluYsrIleYsrProSer 433
QY 1306 ACTACTGTGATGAGCAACAAGCTATGCAAGTCTCCCTGTGAGCAAGCTATGAGAG 1365
DB 434 ThrThrLeuAspGlyThrAsnIleYsrCysGlnPheSerProValGlnIleuSerGlu 453
QY 1366 CAGACTCCCGAAAACAAGCTCTGTAAGAGAGAACCCAGATCCCAAGAGCTGACAGAC 1425
DB 454 GlnThrProGluYsrIleYsrProAspIleYsrGluGluAlaSerIleProIleYsrIleuGluYsr 473
QY 1426 GCCAGGCTTCAAGACGCAAGCAAGCAAGCTTGGATTCGATCAAGACCAAGCAAGTGGC 1485
DB 474 AlaArgProSerAspSerGlnSerIleYsrArgLeuHisSerValArgThrSerSerSerGly 493
QY 1486 ACCGCCAGAGAGTCCCTTATCTCACTGATCAAGAGTGGGCGGAGGAGCAATTAAC 1545
DB 494 ThrAlaGlnArgSerIleuSerProLeuHisArgSerGlySerValGluAspAsnIleYsr 513
QY 1546 CACACCAAGCTTCTTTTGGGCTTTCCACAGCAGCAGACCACTCAAGAGTGTGCTGAC 1605
DB 514 HisThrSerPheLeuPheGlyLeuSerThrSerGlnIleHisIleuThrIleYsrAlaGly 533

```

```

QY 1606 CTGAGGCTTAAAGGCTGAGCACTCGATATCTTGAGCCCCCAAGACCTTACCCCTTCCCTG 1665
DB 534 LeuGluYsrIleuYsrGlyThrPheSerAspIleuAlaProGlnThrSerThrProSerLeu 553
QY 1666 ACCAGCAGCTGTGATTTTCCACAGAGTCTTCACTTCTACTCTGCTCAGCCTTAC 1725
DB 554 ThrSerSerThrPheAlaThrGluSerSerHisPheYsrSerAlaIleYsr 573
QY 1726 GAGAGCAGTCCAGTACTCTGCTTACAGTGCAGCCAGCTGCCCCCTTGGAGAACCA 1785
DB 574 GlyGlySerAlaSerIleYsrAlaYsrCysSerGlnLeuProThrCysGlyAspGln 593
QY 1786 GTCTATTCTGTGCGCAGCGGCGAGAACCAAGTACAGAGCTACTCGCGCGGAGCTGG 1845
DB 594 ValIleSerValArgArgArgGlnYsrProSerAspArgAlaAspSerArgSerThr 613
QY 1846 CATGAAGAGAGCCCTTTGAAAGCAGTTTAAACGAGAAAGCTGCCAAATGAAATTTGGA 1905
DB 614 HisGluGluSerProPheGluYsrGlnPheYsrArgArgSerCysGlnMetGluPheGly 633
QY 1906 GAGAGCATCATGTCAAGAAACAGGTCAAGGAAAGCTGGGAAAGTGGCACTCACTCT 1965
DB 634 GluSerIleMetSerGluAsnArgSerArgGluGluLeuGlyValGlySerGlnSer 653
QY 1966 AGCTTTTGGGCAAGCATGGAATCATTTAGAGTCTCC 2001
DB 654 SerPheSerGlySerMetGluIleIleGluValSer 665

RESULT 14
AAU79162
ID AAU79162 standard; protein; 665 AA.
AC AAU79162;
XX DT
XX 02-JUL-2002 (first entry)
XX DE
XX Human dual-specificity phosphatase-3 (DSP-16) mutant protein #2.
XX KW
XX Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
XX mltogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
XX cancer; graft-versus-host disease; allergy; metabolic disease;
XX abnormal cell growth; abnormal cell proliferation; contact inhibition;
XX cell cycle abnormality; anchorage independent cell growth; apoptosis;
XX intercellular adhesion; DSP-16 modulator; mutant; nuclein.
XX OS
XX Homo sapiens.
XX Synthetic.
XX FH
XX Key Location/Qualifiers
XX FT MISC-difference 244 /note="Wild-type Cys substituted by Ser"
XX PN
XX MO200226997-A2.
XX PD
XX 04-APR-2002.
XX PF
XX 25-SEP-2001; 2001WO-US030124.
XX PR
XX 26-SEP-2000; 2000US-0235487P.
XX PA
XX (CEPT-) CEPTYR INC.
XX PI
XX Luche RM, Wei B;
XX DR
XX MPI; 2002-315802/35.
XX PT
XX New DSP-16 polypeptide, useful for identifying modulators of its
XX activity, which can be used in the treatment of disorders such as
XX Duchenne muscular dystrophy, or cancer.
XX PS
XX Claim 46; Page; 87pp; English.
XX

```

The present invention relates to a new polypeptide, DSP-16, having a 665 amino acid sequence, given in the specification, or a variant having at least 50 % identical residues, which retains the ability to phosphorylate an activated mitogen-activated protein (MAP) kinase. The invention can be used for identifying agents which modulate DSP-16 activity, for modulation of a proliferative response in a cell, survival of a cell, or differentiation of a cell. The cell displays contact inhibition of cell growth or anchorage independent growth and may display altered intercellular adhesion. The agent may modulate apoptosis, or the cell cycle. The identified modulators can be used to treat Duchenne muscular dystrophy, cancer, graft-versus-host disease, autoimmune diseases, allergies, metabolic diseases, abnormal cell growth, abnormal cell proliferation, and cell cycle abnormalities. The present amino acid sequence represents the human dual-specificity phosphatase-3 (DSP-16) mutant protein #2. Note: This sequence is not shown in the specification but is derived from the wild-type human DSP-16 (AAU79156) protein given in figure 2 of the specification

Sequence 665 AA;

Alignment Scores:	
Pred. No.:	0
Score:	421.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	63.12%
DB:	5
	Gaps: 0
	Length: 665
	Matches: 421
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AAU79162 (1-665)

OY		739	TTACTGGGAATCTCCCGCTCGGCACCACTGCTATCGGCTACATCATAGAGGATGAC	798
Db		245	LeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLysArgMetasp	264
OY		799	ATGTCTTAGATGAAGCTTACAGAATTGTGAAAGAAAAAACCTTACTATATTCACAAC	858
Db		265	MetSerLeuAspGluAlaTyrArgPheValLysGluYsarGProThrIleSerProasn	284
OY		859	TTCAATTTCCTGGGCCAATCCTTGACATAAGAAAGAAATTAAAGACCAGCTGGAGCA	918
Db		285	PheasnPheLeuGlyGlnLeuLeuAspTyrGlnLysLysIleLysAsnGlnThrcylala	304
OY		919	TCAGAGGCCAAAGAGCAAATCAAGCGTGTGACCTCGGAGAGCAAGCCAATTAACCTGTCCCT	978
Db		305	SerGlyProLysSerLysLeuLysLeuLeuHisLeuGlnLysProAsnGlnProValPro	324
OY		979	GCTGTCTTCAGAGGGGTGACAGAAAAAGCAGACGCCCCCTCACTCACCTGTGCGCATCT	1038
Db		325	AlaValSerGlnGlyGlyGlnLysSerGlnThrProLeuSerProProCysAlaAspser	344
OY		1039	GCTACCTTCAGAGGACAGAGCAAAAAGCCCGTGCATCCCGCAGCGTSCCAGCGGTGCC	1098
Db		345	AlaThrSerGlnAlaAlaGlyGlnInrProValHisProAlaSerValProSerValPro	364
OY		1099	AGCGTGCAGCCCGTGTGTTAGAGACAGCCCGCTGTACAGCGCGCTCAGTGGCGTGCAC	1158
Db		365	SerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeuSerGlyLeuHis	384
OY		1159	CTGTCCGACAGACGGCTGGAAGACAGCAATAAAGCTCAAGCGTTCTTCTCTGGAATTC	1218
Db		385	LeuSerAlaAspArgLeuGluAspSerAsnLysLeuYsarGSerPheSerLeuAspIle	404
OY		1219	AAATCAGTTTATATTTCAGCCAGCATGGACAGATCTTACATAGGCTTCCTCCGCATCAGAA	1278
Db		405	LysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPheSerSerSerGln	424
OY		1279	GATCGTTTGAATCTACAAACCTTCCACTACTCTGATGGAGCCAAAGCTATGCCAG	1338
Db		425	AspAlaLeuGlnTyrTyrLysProSerThrThrLeuAspGlyThrAsnLysLeuCysGln	444
OY		1339	TTCCTCCCTGTTCACGAATCTAGAGACAGACTCCCAGAAACCAAGTCTGATTAAGAGAA	1398
Db		445	PheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerProAspLysGlnGlu	464

QY	1399	GGCAGACATCCCAAGAAAGCTGCAGACGCGCCAGGCTTTACAGACGCCAGACAGCGATTGG	1458
Db	465	AlAserTleProLysLysLeuGlnTlnRlaArgProSerAspSerGlnSerLysValArgLeu	484
QY	1459	CATTGGCTAGAAACGACGACAGTGGACCGGCCAAGAGTCCCTTTATATCCATCTGCAT	1518
Db	485	HiSerValArgThrSerSerSerGlyThrAlaArgSerLeuLeuSerProLeuHis	504
QY	1519	CGAAGTGGAGCGGTGGAGGACATTTCACACACAGGCTCCTTTCCGCTCCATCCAGCAGC	1578
Db	505	ArgSerLysValAlGlnAspAsnLysTrhSerPheLeuPheGlyLeuSerThrSer	524
QY	1579	CACACAGACCTTCACGAAAGTGTGGCTGGGCTTGAGGCTGGCACTCGGATATCTTG	1638
Db	525	GlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTrpHisSerAspIleLeu	544
QY	1639	GCCCCCAAGACTCTTACCCCTTCCTGACACGACGCTGGTATTTGGCACAGAGTCTCTCA	1698
Db	545	AlaProGlnThrSerThrProSerLeuThrSerSerTrpLysPheAlaThrGlnSerSer	564
QY	1699	CACCTTCACCTCGCTCAGCAGCATCTCAAGAGGAGGACGACCGACAGTACTCGCTCAGCTGC	1758
Db	565	HisPheLysSerLysSerAlaIleLysTrpGlyLysSerAlaSerLysSerAlaLysSerCys	584
QY	1759	AGCCAGCTGCCCACTTGGCGAGACCAAGTCTATCTGTGTGGCAGCGCGCAAAAGCCAACT	1818
Db	585	SerGlnLeuProThrLysGlyLysArgGlnValLysSerAlaArgLysArgGlnLysProSer	604
QY	1819	GACAGACCTGACTCTGCGCGCGAGCTGCATGAAGAGAGCCCTTTGAAAGCAGTTTAA	1878
Db	605	AspArgLysAlaAspSerLysArgSerTrpHisGlnGlnSerProPheGlnLysGlnPheLys	624
QY	1879	CGCAGAAAGCTGCCAAATGGAATTTGGAGAGAGCATCTATCTCAGAGAAACAGTACCGGAA	1938
Db	625	ArgArgSerCysGlnMetGlnPheGlyLysSerIleMetSerGlnAsnArgSerArgGln	644
QY	1939	GAGCTGGGAAAGTGGGAGCTCAAGTCTTAAGCTTTTGGGACAGCATGAAATCATTTAGAGTC	1998
Db	645	GlnMetLysLysValGlySerGlnSerSerPheSerGlySerMetGlnIleIleGlnVal	664
QY	1999	TCC 2001	
Db	665	Ser 665	
RESULT 15			
ABR52352			
ID	ABR52352 standard; protein: 665 AA.		
XX	ABR52352;		
AC			
DT	19-JUN-2003 (first entry)		
DB	Protein relating to the invention SEQ ID NO: 42.		
XX	antiproliferative; hepatotropic; nephrotropic; antiarthritic;		
KW	antiproliferative; cardiant; cytostatic; gene therapy; liver disease;		
KW	proliferative disorder; renal failure; cardiovascular disorder;		
KW	immunological disorder; arthritis; psoriasis; congenital heart defect;		
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.		
XX	Homo sapiens.		
OS	Homo sapiens.		
PN	MO200257460-A2.		
XX	25-JUL-2002.		
PF	20-DEC-2001; 2001WO-US050459.		
XX	20-DEC-2000; 2000US-0256868P.		
PR	30-MAR-2001; 2001US-0280186P.		
PR	01-MAY-2001; 2001US-0287735P.		
PR	05-JUN-2001; 2001US-0295848P.		

FT	Modified-site	/note= "O-phosphorylated"	82
FT		/note= "O-phosphorylated"	85. .298
FT	Region	/note= "H1-type dual specificity phosphatase signature"	91
FT	Modified-site	/note= "O-phosphorylated"	190
FT	Modified-site	/note= "N-glycosylated"	212
FT	Modified-site	/note= "N-glycosylated"	214
FT	Active-site	/note= "O-phosphorylated"	220. .280
FT		/note= "tyrosine specific protein phosphatase signature"	237. .278
FT	Region	/note= "Y phosphatase signature"	266
FT	Modified-site	/note= "O-phosphorylated"	280
FT	Modified-site	/note= "O-phosphorylated"	300
FT	Modified-site	/note= "N-glycosylated"	369
FT	Modified-site	/note= "O-phosphorylated"	393
FT	Modified-site	/note= "O-phosphorylated"	421
FT	Modified-site	/note= "O-phosphorylated"	422
FT	Modified-site	/note= "O-phosphorylated"	434
FT	Modified-site	/note= "O-phosphorylated"	439
FT	Modified-site	/note= "O-phosphorylated"	468
FT	Modified-site	/note= "O-phosphorylated"	471
FT	Modified-site	/note= "O-phosphorylated"	479
FT	Modified-site	/note= "O-phosphorylated"	528
FT	Modified-site	/note= "O-phosphorylated"	590
FT	Modified-site	/note= "O-phosphorylated"	597
FT	Modified-site	/note= "O-phosphorylated"	605
FT	Modified-site	/note= "O-phosphorylated"	610
FT	Modified-site	/note= "O-phosphorylated"	613
FT	Modified-site	/note= "O-phosphorylated"	618
FT	Modified-site	/note= "O-phosphorylated"	628
FT	Modified-site	/note= "O-phosphorylated"	641
FT	Modified-site	/note= "N-glycosylated"	643
FT	Modified-site	/note= "O-phosphorylated"	
XX			
PN	W0200120004-A2.		
PD	22-MAR-2001.		
XX			
PP	14-SEP-2000; 2000WO-US025515.		
XX	15-SEP-1999; 99US-0154141P.		
PR			
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
PI	Yue H, Tang YT, Bandman O, Hillman JI, Baughn MR, Azimzai Y,		

[illegible]

Db		354	ProValHisProAlaSerValProSerValGlnProSerLeuLeuGlnAsp	373
Qy		1126	AGCCCGCTGGTACAGCGCGCTGATGCGACTGTCCGACAGACAGGCTGAAAGACGC	1185
Db		374	SerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSer	393
Qy		1186	AATAAGCTCAACGCTTCTCTCTCTGGAATATCAAAATGATTTTCATATTCAGCCACGATG	1245
Db		394	AsnIysLeuLysArgSerPheSerLeuAspIleIysSerValSerTyrSerAlaSerMet	413
Qy		1246	GCAGATCCCTTACAGAGGCTCTCCCTCCACAGAAAGATGCTTTGGAATACTACAAACCTTCC	1305
Db		414	AlaAlaSerLeuHisGlyPheSerSerSerGlnuSpAlaLeuGluTyrTyrLeuProSer	433
Qy		1306	ACTACTCTGATGGAGCCACCAACAGCTATGCCAGTTCTCCCTGTTCAGAACTATCGAG	1365
Db		434	ThrThrLeuAspGlyThrAsnIysLeuGlnPheSerProValGlnGlnLeuSerGln	453
Qy		1366	CAGACTCCCGAAACCAATC--CTGATNAGAGAAAGCCAGCATCCCAAGAAAGCTGCAGA	1423
Db		454	GlnThrProGlnThrSerSerLeuIleArgIysProAlaSerProArgSerCysArg	473
Qy		1424	CCGCCAGGCTTCAGACAGCCAGACGACATTCGATTCGGTACGAAACAGACGACGTG	1483
Db		474	ProProGlyLeuGlnThrAlaArgAlaSerAspCysIleArgSerGlnProAlaAlaVal	493
Qy		1484	GCACCCGCCACAGAGTCCCTTTATCTCCACTGCATCGAAGTGGAGCGCTGAGAGCAATT	1543
Db		484	AlaProProAlaGlyProPheIysLeuHisCysIleGluValGlyAlaThrArgThrIle	513
Qy		1544	ACCACACACAGCTTCTTTTGGCCCTTTGCCACACAGCAGACACACCTCAGAAAGTCTGCTG	1603
Db		514	ThrThrProAlaSerPheSerAlaPheProProAlaSerSerThrSerArgSerLeuLeu	533
Qy		1604	GCCTTGGC--CTTAAGGCGTGGCATCTCGGATATCTTGGGCCCCCAGACCTCTAACCCCTTCC	1662
Db		534	AlaThrAlaLeuLysGlyTyrPheHisSerAspIleLeuAlaProGlnThrSerThrProSer	553
Qy		1663	CTGACACACAGCTGGTATTTTGGCCACAGACAGTCCCTCAGCTTCTGCTGCTCAGCCATC	1722
Db		554	LeuThrSerSerThrTyrPheAlaThrGlnSerSerHisPheTyrSerAlaSerAlaIle	573
Qy		1723	TACGAGGACAGTGCAGGTTACTCTGCTCTACAGCTGACGACGCTGCCACTTGGCGGAGAC	1782
Db		574	TyrGlyGlySerAlaSerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAsp	593
Qy		1783	CAAGTCTATTTCTGTCCGACGGCGGAGAAAGCCAAAGTACAGAGCTGACTCGCGGCGAGGC	1842
Db		594	GlnValTyrSerValArgArgGlnIysProSerAspArgAlaAspSerArgArgSer	613
Qy		1843	TGGCATGAAAGAGAGCCCTTTGAAAGAGAGTTTAAAGCAGACAGCTGCCAAATGGAATTT	1902
Db		614	TrpHisGlnGlnSerProPheGlnIysGlnPheLysAspArgSerCysGlnMetGlnPhe	633
Qy		1903	GGAGAGACATCATCTCAGAGAAACAGTCCAGCGGAAAGAGCTGGGGAAGTGGGAGTCTAG	1962
Db		634	GlyIysSerIleMetSerGlnAsnArgSerArgGlnGlnLeuGlyLysValGlySerGln	653
Qy		1963	TCTAGCTTTTGGGCGACGATGGAAATCATTTGAGGTCTCC	2001
Db		654	SerSerPheSerGlySerMetGlnIleIleGlnValSer	666
RESULT 17				
ABR52424				
ID	ABR52424 standard: protein; 664 AA.			
XX	ABR52424;			
XX	19-JUN-2003 (first entry)			
DE	Protein relating to the invention SEQ ID NO: 190.			

KW	antiproliferative; hepatotropic; nephrotropic; antiarthritic;
KV	antiproliferative; cardiact; cytotactic; gene therapy; liver disease;
KM	proliferative disorder; renal failure; cardiovascular disorder;
KW	immunological disorder; arthritis; psoriasis; congenital heart defect;
KM	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX	
OS	Homo sapiens.
XX	
FN	WO200257460-A2.
PD	
XX	
XX	25-JUN-2002.
PE	
XX	20-DEC-2001; 2001WO-US050459.
PR	
XX	20-DEC-2000; 2000US-0256868P.
PR	30-MAR-2001; 2001US-0280186P.
PR	01-MAY-2001; 2001US-0287735P.
PR	05-JUN-2001; 2001US-0295848P.
PR	25-JUN-2001; 2001US-0300465P.
XX	
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.
PJ	
PI	Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
PI	Siemens N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
PI	Krysstek S, Mcatee P, Suchard S, Banas D;
DR	WPI; 2002-599721/64.
PT	
PT	Novel polynucleotides encoding human phosphatase polypeptides useful in
PT	the prevention or treatment of e.g. proliferative and cardiovascular
PT	disorders.
XX	
PS	Example 57; Page 500-501; 801pp; English.
CC	The invention relates to a novel isolated nucleic acid comprising a
CC	polynucleotide having a nucleotide sequence selected from 40
CC	polynucleotides fully defined in the specification. The polynucleotide of
CC	the invention has antiproliferative, hepatotropic, nephrotropic, CC
CC	antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
CC	polypeptide may have a use in gene therapy. A polynucleotide or
CC	polypeptide of the invention is useful for preventing, treating or
CC	ameliorating a medical condition, e.g. a proliferative disorder. They are
CC	also useful for treating e.g. liver disease, renal failure, immunological
CC	disorders including arthritis and psoriasis, cardiovascular disorders
CC	such as congenital heart defects and congestive heart failure, and
CC	cancer. A method of the invention is useful for diagnosing a pathological
CC	condition or susceptibility to a condition in a subject. The present
CC	sequence is used in the exemplification of the invention
XX	
SQ	Sequence 664 AA:
Alignment Scores:	
Pred. No. :	7.26e-267 Length: 664
Score:	281.00 Matches: 281
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	42.13% Indels: 0
DB:	5 Gaps: 0
US-10-029-345A-108 COPY_532_2532 (1-2001) x ABR52424 (1-664)	
OY	913 GGAGCATCAGGCGCAAGAAGCAAACTCAAGCTGTGCATCTGGAGAACCAATGAACCT 972
DB	302 GtTlaaSerGtLPrOlySserLyLeuLysleuLysleuHtsleuGtLylPrOaBnGlUpRo 321
OY	973 GTCCCTGCTGTCTCAGAGGCTGACAGAAAAGCAAGCGCCCTCAGTCACCTGTGCC 1032
DB	322 ValProAlaValSerGIuGIyGLyGIlnlysSerGIuThrProLeuSerProPCyala 341
OY	1033 GACTCTGTAACCTAGAGGCGACGAGCAAAAGCCCGTGATCCCGCAGCGTCCCAAG 1092
DB	342 AsperalathrTSerGIuaAlaAGLyGIlnmrGrProVALHisProAlaSerValProSer 361

QY 1093 GTGCCAGCCGTGACCCGTCGTTAGAGACAGACCCGCTGTACAGAGCGCTCAGTGG 1152
|
|
|
Db 362 ValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeuSerGly 381
|
|
|
QY 1153 CTGACCTGTCCGACAGACAGCTGTGAAGACAGCAATTAAGCTCAAGCCTTCTCTCTG 1212
|
|
|
Db 382 LeuHISLeuSerAlaAspArgLeuGlnAspSerAsnGlySerValArgSerPheSerLeu 401
|
|
|
QY 1213 GATATCAATCAGTTTCAATATATTCAGCCAGATGGCAGATCCCTTAACATGGCTTCTCTCA 1272
|
|
|
Db 402 AspiLeuSerValSerTyrSerAlaSerMetAlaAlaSerLeuHISGlyPheSerSer 421
|
|
|
QY 1273 TCAGAAATGCTTTGGAATATACAAACCTTCCACTCTGTGATGGACCAACAGCTA 1332
|
|
|
Db 422 SerGlnAspAlaLeuGlnTyrTyrTyrProSerThrThrLeuAspGlyThrAsnGlyLeu 441
|
|
|
QY 1333 TGCACGTTTCCCTGTTTCAGAACTATCGAGAGAGACTCCGAAACAGTCTGTATAG 1392
|
|
|
Db 442 CysGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerProAspGly 461
|
|
|
QY 1393 GAGGAGCCAGCATCCCAAGAGCTGACAGCCGCGCTTCAGACAGCCAGAGCAAG 1452
|
|
|
Db 462 GlnGlnAlaSerLeuProValGlySerGlnThrAlaArgProSerAspSerGlnSerLys 481
|
|
|
QY 1453 CGATTGCAATTCGTACAGACCCAGAGAGTGGACCCGCAAGTCCCTTTATCTCCA 1512
|
|
|
Db 482 ArgLeuHISSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeuSerPro 501
|
|
|
QY 1513 CTGATGAGAGTGGGACCGTGGAGAGCAATTACCAACCGCTTCTTTCGCGCTTTTC 1572
|
|
|
Db 502 LeuHISArgSerGlySerValGlnAspAsnTyrHISThrSerPheLeuPheGlyLeuSer 521
|
|
|
QY 1573 ACCAGCCAGAGACACTCAGAAAGTGTGCTGCGCTGGGCTTAAAGGCTGACACTGGAT 1632
|
|
|
Db 522 ThrSerGlnGlnHISLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheISerAsp 541
|
|
|
QY 1633 ATCTTGGCCCCCAGACCTTACCCCTTCCCTTACCAAGACAGCTGATTTTGGCACAG 1692
|
|
|
Db 542 IleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPyrPheAlaThrGln 561
|
|
|
QY 1693 TCCCTCACACTTACTCTGCTCAGACCATAGAGAGGAGCGCCAGTTACTCTGCTAC 1752
|
|
|
Db 562 SerSerHISpheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSerAlaTyr 581
|
|
|
QY 1753 AGC 1755
|
|
|
Db 582 Ser 582
|
|
|
RESULT 18
AAM25744
ID AAM25744 standard; protein; 672 AA.
AC AAM25744;
XX
XX
DT 16-OCT-2001 (first entry)
XX
XX
DE Human protein sequence SEQ ID NO:1259.
XX
XX
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antithyroid; immunosuppressive;
KW antibacterial; endocrine; cardiac; central nervous system; vitruclide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnery; antidiabetic; osteopathic; eczema;
KW dermatological; antiallergic; antisthmatic; antiparkinsonian; infection;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmune;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.

XX
OS Homo sapiens.
XX
PN MO200153455-A2.
XX
PD 26-JUL-2001.
XX
PF 22-DEC-2000; 2000MO-US035017.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
XX
XX (HYSR-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR MPI; 2001-457603/49.
XX
DR N-PSDB; AAH99685.
XX
PT Isolated human polynucleotides encoding polypeptides, useful for the
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
PS Claim 20; Page 260; 1217p; English.

AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antithyroid; immunosuppressive; antibacterial; endocrine; cardiac;
CC central nervous system; vitruclide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
CC antidiabetic; osteopathic; dermatological; antiallergic; antisthmatic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders

XX
SQ Sequence 672 AA;

Alignment Scores:
Pred. No.: 3,02e-244 Length: 672
Score: 258.00 Matches: 258
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.68% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AAM25744 (1-672)

QY 586 ACCGTGCCAAGGCTGACCTTATCCCGAGTCTGATTTCTGCGTGGCTGTGAATGAC 645
|
|
|
Db 201 ThrCysProLysProAspPheIleProGlnSerHISpheLeuArgValProValAsnAsp 220
|
|
|
QY 646 AGCTTTGTAGAAATTTTGGCCGTGTGACAAATCAGTATGATTTTCAATTGGAAGAACA 705
|
|
|
Db 221 SerPheCysGlnLysIleLeuProTyrPheAspLysSerValAspPheIleGlnLysAla 240
|
|
|
QY 706 AAGCCTCCAAATGATGTGTTTGTAGTCACTGTTTACCTGGAGATCCGCTCGCACACC 765
|
|
|
Db 241 LysAlaSerIleGlnGlyCysValIleValHISCysLeuAlaIleGlyIleSerHISerAlaThr 260
|
|
|
QY 766 ATGCGTATGCGCTTACATCATGAAGAGATGACATGCTTTTATGAAAGCTTAACAGATT 825
|
|
|

Dd		261	lealaillelaIatYrilleMeLysarginehlsphmetserLeuaAspGluAlaTyArgphe	280
Oy		826	GTGAAGAAGAAAAGACCTACTATATCTCCAAATCTCAATTTTTCGGGCCAACCCTCGAC	885
Dd		281	VallysglulysargProthrilleserProasnpheanphenleuglyglnleuluasp	300
Oy		886	TATAGAAGAAGATTAAAGACCAAGCTGGAGCATGAGGCGCAAAGACAATCAGACTG	945
Dd		301	TyrGlulysylleLysasnGlnThrslglalaserglyprolysserLysleuluysleu	320
Oy		946	CTGCACCTCGAGAAAGCAAATGAACCTTGCTCCCTGCTGCTCAGAGAGGTGACAGAAAAAGC	1005
Dd		321	leuhisleuGluLysProlasnglunproValProalaValserglunglyglnhysser	340
Oy		1006	GAGACGCCCTTCAGTCCACCTCTGGCCGACTCTGCTACCTCAGAGGCACAGAACAAAGG	1065
Dd		341	GluThrProlseurProProcysalaAspserralatrnserrglualalaglyGlnarg	360
Oy		1066	CCCCGGCATCCCGGACGGGTGGCCCGCCGCGCCAGCGCGTGGCTTTAGAGAC	1125
Dd		361	ProvalHisProlaserValProserValProserValGlnProserLeuluengluasp	380
Oy		1126	AGCCCGCTGGTAGACAGCGCTCAGTGGGTGACCGCTGCCGAGACAGCGCTGGAAGACAGC	1185
Dd		381	SerProleuValGlnAlaleuserylLeuhlsbeuseralalapargleuGluhsper	400
Oy		1186	AATAAGCTCAAGCGTTCCTCTCTCTGGATATCAATCAAGTTTATTCAGCCAGCATG	1245
Dd		401	AsnlysleulySargSerPhesterleuaspIleuyservalserTyserAlaserMet	420
Oy		1246	GCAGATCCTTAACATGAGCTTCTCTCATCAGAAAGTGGTTTGAATACTACAAACCTTCC	1305
Dd		421	AlaAlaserLeuHlglyPheSerSerSerGluaspAlaleuGluTy-TyrlsPserSer	440
Oy		1306	ACTACTCGGAATGGAGCAACCAAGCATGCCAGTCTGCCCGTTCAGAGAACTA	1359
Dd		441	ThrnTrtleuaspglyThrasnLysleucysglnPheSerProvalGlngluenu	458
RESULT 19				
ID	ABR52385			
ID	ABR52385	standard; protein; 660 AA.		
XX	ABR52385;			
DT	19-JUN-2003	(first entry)		
DE	Protein relating to the invention SEQ ID NO: 114.			
KW	antiproliferative; hepatotropic; nephrotropic; antiarthritic;			
KM	antiapoptotic; cardiant; cytosstatic; gene therapy; liver disease;			
KW	proliferative disorder; renal failure; cardiovascular disorder;			
KM	immunological disorder; arthritis; psoriasis; congenital heart defect;			
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.			
XX	Mus musculus.			
OS				
PN	WO200257460-A2.			
XX				
PD	25-JUL-2002.			
XX				
PF	20-DEC-2001; 2001WO-US050459.			
XX				
PR	20-DEC-2000; 2000US-0256868P.			
PR	30-MAR-2001; 2001US-0280186P.			
PR	01-MAY-2001; 2001US-0287735P.			
PR	05-JUN-2001; 2001US-0295848P.			
PR	25-JUN-2001; 2001US-0300465P.			
XX				
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.			
PI	Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,			
PI	Slennre N, Bol D, Schieven G, Finger U, Todderud CG, Bassolino D,			
PI	Krystek K, Mattee P, Suchard S, Baanae D;			

XX	WP1: 2002-599721/64.
DR	N-P8DB; ACC60560.
XX	
PT	Novel polynucleotides encoding human phosphatase polypeptides useful in
PT	the prevention or treatment of e.g. proliferative and cardiovascular
PT	disorders.
XX	
PS	Disclosure: fig 12; 801pp; English.
XX	
XX	The invention relates to a novel isolated nucleic acid comprising a
CC	polynucleotide having a nucleotide sequence selected from 40
CC	polynucleotides fully defined in the specification. The polynucleotide of
CC	the invention has antiproliferative, hepatotropic, nephrotropic, the
CC	antiarthritic, antiposoriatic, cardiant, and cytostatic activity. The
CC	polynucleotide may have a use in gene therapy. A polynucleotide or
CC	polypeptide of the invention is useful for preventing, treating or
CC	ameliorating a medical condition, e.g. a proliferative disorder. They are
CC	also useful for treating e.g. liver disease, renal failure, immunological
CC	disorders including arthritis and psoriasis, cardiovascular disorders
CC	such as congenital heart defects and congestive heart failure, and
CC	cancer. A method of the invention is useful for diagnosing a pathological
CC	condition or susceptibility to a condition in a subject. The present
CC	sequence is used in the exemplification of the invention
XX	
SQ	Sequence 660 AA;
	Alignment Scores:
	Pred. No.: 1,876-53 Length: 660
	Score: 64.00 Matches: 64
	Percent Similarity: 100.00% Conservative: 0
	Best Local Similarity: 100.00% Mismatches: 0
	Query Match: 9.60% Indels: 0
	DB: 5 Gaps: 0
US-10-023-345A-108_COPY_532_2532 (1-2001) x ABR52385 (1-660)	
OY	346 AGCTTCAACTCTGTTACCTGCTTGACAGGTGGAGTTGCTGAGTCTCTGTTGTTCCCT 405
DB	114 SerPheAaSerValHisIleuEuAaIaGlyGlyPheAlaGluPheSerArgCySPhePro 133
OY	406 GGCCTCTGTGAAGAAATCCACTAGTCCTACTGCAATTCCTACGCTTGCTTAACT 465
DB	134 GlyLeuCySGluGlyLysSerThrIleuValProThrCyAlaSerGlnProCyValuPro 153
OY	466 GTTGCCACATTTGGGCCCAACCCGAATTCTTCCCAATCTTATCTTGCGTGCACGAGAT 525
DB	154 ValAlaAaenIIeGlyProThrArgIleuProAaenLeuTrIleuGlyCySGlnAaGAsp 173
OY	526 GTCTCAACAAG 537
DB	174 ValIeuAaenLys 177
RESULT 20	
ID	ADE08458 standard; protein; 579 AA.
XX	ADE08458;
XX	
AC	29-JAN-2004 (first entry)
DT	
XX	
DE	Novel protein (useful for identifying genetic disorders) #613.
XX	
XX	novel gene; novel protein; tissue marker; molecular weight marker;
KW	chromosome marker; genetic disorder.
XX	
OS	Unidentified.
XX	
XX	WO2003054152-A2.
PN	
XX	03-JUL-2003.
PD	
XX	
PF	10-DEC-2002; 2002WO-US039555.

XX 10-DEC-2001; 2001US-0339739P.
PR 11-DEC-2001; 2001US-0339453P.
PR 14-MAR-2002; 2002US-0365091P.
PR 14-MAR-2002; 2002US-0365384P.
PR 12-APR-2002; 2002US-0372315P.
PR 12-APR-2002; 2002US-0372615P.
PR 22-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Asundi V, Goodrich RW, Ren P, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX
DR WPI: 2003-569235/53.
DR N-PSDB; ADE07547.
XX
PT New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX
PS Claim 20; SEQ ID NO 1524; 1177bp; English.
XX
CC The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present amino acid sequence represents a protein
CC of the invention.
XX
SQ Sequence 579 AA;
XX
Alignment Scores:
Pred. No.: 3.57e-10 Length: 579
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 7 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x ADE08458 (1-579)
QY 730 GTGCAGTGTAGGATGATCCGCTCCGACCATGCGCTATCGCTACATCATGANG 789
Db 186 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetCys 205
RESULT 21
ABR52382
ID ABR52382 standard; protein; 625 AA.
XX
AC ABR52382;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein relating to the invention SEQ ID NO: 110.
XX
KW anti-proliferative; hepatotropic; nephrotropic; antiarthritic;
KW anti-ischaemic; cardiatic; cytosolic; gene therapy; liver disease;
KW proliferative disorder; renal failure; cardiovascular disorder;
KW immunological disorder; arthritis; psoriasis; congenital heart defect;
KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
OS Homo sapiens.
XX
PN WO200257460-A2.
XX
PD 25-JUL-2002.
XX
PF 20-DEC-2001; 2001WO-US050459.

XX 20-DEC-2000; 2000US-0256688P.
PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
PI Siemsen N, Bol D, Schieven G, Finger J, Todderud CG, Baasolino D;
PI Kysceck S, Mcatee P, Suchard S, Banas D;
XX
DR WPI: 2002-599721/64.
XX
PT Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
PS disclosure; Fig 14; 801pp; English.
XX
CC The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has anti-proliferative, hepatotropic, nephrotropic,
CC antiarthritic, anti-ischaemic, cardiatic, and cytosolic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX
SQ Sequence 625 AA;
XX
Alignment Scores:
Pred. No.: 3.53e-10 Length: 625
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 5 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x ABR52382 (1-625)
QY 730 GTGCAGTGTAGGATGATCCGCTCCGACCATGCGCTATCGCTACATCATGANG 789
Db 244 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetCys 263
RESULT 22
ABR52350
ID ABR52350 standard; protein; 625 AA.
XX
AC ABR52350;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein relating to the invention SEQ ID NO: 39.
XX
KW anti-proliferative; hepatotropic; nephrotropic; antiarthritic;
KW anti-ischaemic; cardiatic; cytosolic; gene therapy; liver disease;
KW proliferative disorder; renal failure; cardiovascular disorder;
KW immunological disorder; arthritis; psoriasis; congenital heart defect;
KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
OS Homo sapiens.
XX
PN WO200257460-A2.
XX
PD 25-JUL-2002.

XX 20-DEC-2001; 2001WO-05050459.
PF
XX
PR 20-DEC-2000; 2000US-0256868P.
PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Jackson DG, Feder J, Nelson T, Mintler G, Ramanathan C, Lee L,
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Basolino D,
PI Krytek S, Macree P, Suchard S, Banas D;
XX
XX WPI; 2002-599721/64.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
XX
XX Disclosure; Fig 10; 801pp; English.
XX
XX The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antidiabetic, antiproliferative, cardiac, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX
XX Sequence 625 AA;
SQ
Alignment Scores:
Pred. No.: 3.53e-10 Length: 625
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x ABR52350 (1-625)
QY 730 GTGCACTGTTTGGGATCTCCGCTCCGCCACCATGCTATCGCCTACATGAG 789
DB 244 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 263
RESULT 23
ABG73440
ID ABG73440 standard; protein; 625 AA.
XX
XX ABG73440;
AC
XX
XX 10-MAY-2003 (first entry)
DT
XX
XX Human dual specific phosphatase 8 polypeptide.
DE
XX
XX Human; dual specific phosphatase 8; enzyme; infection; inflammation;
KW tumour formation; cytostatic; antiinflammatory.
XX
XX Homo sapiens.
OS
XX US6482644-B1.
FN
XX 19-NOV-2002.
PD
XX

PF 01-AUG-2001; 2001US-00920668.
XX
XX 01-AUG-2001; 2001US-00920668.
PR
XX
XX (ISIS-) ISIS PHARM INC.
PA
XX
XX Cowse LM;
PI
XX
XX WPI; 2003-298140/29.
DR N-PSDB; ABX10760.
XX
XX
XX New antisense compound targeted to a nucleic acid encoding human dual
PT specific phosphatase 8, for modulating gene expression and treating
PT diseases associated with expression of the phosphatase in humans.
XX
XX
XX Disclosure; Col 49-54; 36pp; English.
XX
XX The invention relates to a compound targeted to the coding region of a
CC nucleic acid encoding human dual specific phosphatase 8, where the
CC compound specifically hybridises with the region and inhibits the
CC expression of human dual specific phosphatase 8. The compound is useful
CC for inhibiting the expression of human dual specific phosphatase 8 in
CC cells or tissues, and for treating an animal, particularly a human,
CC suspected of having or being prone to a disease or condition associated
CC with expression of dual specific phosphatase 8. The compound is useful
CC for diagnosis, therapeutics and as a research reagent, e.g. to prevent
CC or delay infection, inflammation or tumour formation, and to distinguish
CC between functions of various members of a biological pathway. This
CC sequence represents human dual specific phosphatase 8
XX
XX Sequence 625 AA;
SQ
Alignment Scores:
Pred. No.: 3.53e-10 Length: 625
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x ABG73440 (1-625)
QY 730 GTGCACTGTTTGGGATCTCCGCTCCGCCACCATGCTATCGCCTACATGAG 789
DB 244 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 263
RESULT 24
AAM29150
ID AAM29150 standard; protein; 663 AA.
XX
XX AAM29150;
AC
XX
XX 15-DEC-1997 (first entry)
DT
XX
XX Dual-specific murine threonine-tyrosine phosphatase M3/6.
DE
XX
XX murine; phosphatase; M3/6; dual specificity; Threonine; Tyrosine;
KW inactivate; mitogen activated protein kinase; MAP-K; cdcs25 pT; yeast;
KW trinucleotide repeat; repeat expansion; neurodegenerative; proliferative;
KW diagnosis; tumour; lung; brain; chromosomal deletion.
XX
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 29..49
FT Misc-difference 117..136
FT Misc-difference 117..136
FT /note="region of homology with yeast cdcs25"
FT Domain 244..253
FT /label=catalytic_domain
XX MO9706245-A1.
XX

PD	27-SEP-2001.				
XX					
PF	23-MAR-2001; 2001WO-US009231.				
PR	23-MAR-2000; 2000US-0191637P.				
PR	11-JUL-2000; 2000US-00614150.				
XX					
PA	(PEKE) PE CORP NY.				
XX					
PI	Venter JC, Adams M, Li PWD, Myers EW;				
DR	WPI; 2001-656860/75.				
DR	N-PSDB; ABL07630.				
XX					
PT	New isolated nucleic acid detection reagent for detecting 1000 or more				
PT	genes from Drosophila and for elucidating cell signaling and cell-cell				
XX	interactions.				
PS	Disclosure; SEQ ID NO 17373; 21bp + Sequence Listing; English.				
XX					
CC	The invention relates to an isolated nucleic acid detection reagent				
CC	capable of detecting 1000 or more genes from Drosophila. The invention is				
CC	useful in developmental biology and in elucidating cell signaling and				
CC	cell-cell interactions in higher eukaryotes for the development of				
CC	insecticides, therapeutics and pharmaceutical drugs. The invention				
CC	discloses genomic DNA sequences (AB16176-AB130511), expressed DNA				
CC	sequences (AB101840-AB16175) and the encoded proteins (AB57737-				
CC	AB572072). The sequence data for this patent did not form part of the				
CC	printed specification, but was obtained in electronic format directly				
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences				
XX					
SO	Sequence 476 AA;				
	Alignment Scores:				
	Pred. No.:	0.00281	Length:	476	
	Score:	13.00	Matches:	13	
	Percent Similarity:	100.00%	Conservative:	0	
	Best Local Similarity:	100.00%	Mismatches:	0	
	Query Match:	1.95%	Indels:	0	
	DB:	4	Gaps:	0	
US-10-029-345A-108_COPY_532_2532 (1-2001) x ABB63527 (1-476)					
OY	742 GCTGGATCTCCCGCTCCGCCACCATCGCTATCGCCTAC	780			
DB	218 AlAGIyIleSerArGserAlaThrIleAlaIleAlaIyr	230			
RESULT 27					
ID	ABG00724 standard; protein; 836 AA.				
XX					
AC	ABG00724;				
XX					
DT	13-FEB-2002 (first entry)				
XX					
DE	Novel human diagnostic protein #715.				
XX					
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;				
XX	food supplement; medical imaging; diagnostic; genetic disorder.				
OS	Homo sapiens.				
XX					
PN	WO200175067-A2.				
XX					
PD	11-OCT-2001.				
XX					
PF	30-MAR-2001; 2001WO-US008631.				
XX					
PR	31-MAR-2000; 2000US-00540217.				
XX					
PR	23-AUG-2000; 2000US-00649167.				
XX					
PA	(HYSE-) HYSEQ INC.				
XX					

PI	Drimanac T., Liu C., Tang Y.T.
XX	
DR	WPI; 2001-639362/73.
DR	N-PSSB; AAS64911.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
PS	Claim 20; SEQ ID NO 31083; 103bp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC	amino acid sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 836 AA:
	Alignment Scores:
	Pred. No.: 0.0251 Length: 836
	Score: 12.00 Matches: 12
	Percent Similarity: 100.00% Conservative: 0
	Best Local Similarity: 100.00% Mismatches: 0
	Query Match: 1.80% Indels: 0
	DB: Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x ABG00724 (1-836)	
OY	724 GTTCTAGTGACGTGTAGCTGGATCTCCGCCTCC 759
DB	744 ValLeuValHisCysLeuAlaIleLysSerArgSer 755
RESULT 28	
ID	ABU29344 standard; protein; 498 AA.
AC	ABU29344;
XX	
DT	19-JUN-2003 (first entry)
XX	
DE	Protein encoded by prokaryotic essential gene #14871.
KX	Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX	
OS	Enterococcus faecalis.
XX	
PN	WO200277183-A2.
PD	
FD	03-OCT-2002.
PF	
PP	21-MAR-2002; 2002WO-US009107.
XX	
PR	21-MAR-2001; 2001US-0081524Z.
RR	06-SEP-2001; 2001US-00948993.
FR	25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zykkind JW;
PI Wall D, Trivick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA33214.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 57268; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 498 AA;
XX
Alignment Scores:
Pred. No.: 2 49
Score: 10.00 Length: 498
Percent Similarity: 100.00% Matches: 10
Best Local Similarity: 100.00% Conservative: 0
Query Match: 1.55% Mismatches: 0
DB: Indels: 0
Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x ABU29344 (1-498)
OY 103 GCGGCTATCATTTAGCAGACTTTTTCG 74
DB 342 AAtAGlyTyrGlnLeuAlaIleuPhePro 351
RESULT 29
ID ADA44807
XX ADA44807 standard; protein; 473 AA.
AC
XX ADA44807;
XX 04-DEC-2003 (first entry)
DT
XX

DE CD4/TCR CD3epsilon chain chimeric protein CD4epsilon1015, SEQ ID NO:2.
XX
XX HIV-1 infection; human immunodeficiency virus-1; CD4+ cell; chimeric CD4;
KW endoplasmic reticulum; ER retention; envelope protein gp160;
KW T cell receptor CD3epsilon chain; C-terminal domain; CD4epsilon1015;
KW gene therapy; human; receptor.
XX
OS Chimeric.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..458
FT /label= CD4
FT Region 459..473
FT /note= "Part of the C-terminal domain of the T cell
FT receptor CD3epsilon chain"
XX
XX WO2003076468-A1.
XX
XX 18-SEP-2003.
XX
XX 14-MAR-2003; 2003WO-ES000120.
XX
XX 14-MAR-2002; 2002ES-00000616.
XX
XX (CNSG) CONSENSO SUPERIOR INVESTIGACIONES CIENTIF.
XX
XX Alarcon Sanchez BJ, San Jose Martinez ME, Zaldivar Notario I;
XX Gomez Buendia M;
XX WPI; 2003-779059/73.
XX N-PSDB; ADA44806.
XX
PT Composition for treating or preventing human immune deficiency virus,
PT comprises CD4 chimeric protein having a protective effect in trans, or
PT related nucleic acid.
XX
XX Claim 5; Page 33-35; 43pp; Spanish.
XX
CC The invention relates to a composition for the treatment or prevention of
CC human immunodeficiency virus-1 (HIV-1) infection. The composition
CC comprises CD4+ cells that have been transduced with a vector that encodes
CC a chimeric CD4 molecule which is capable of being retained in the
CC endoplasmic reticulum (ER). The invention also encompasses the use of a
CC soluble protein factor produced by CD4+ cells that have been transduced
CC with a vector encoding a chimeric CD4 protein; and the use of an
CC expression system encoding a chimeric CD4 protein; and the use of an
CC chimeric CD4 molecule binds to the HIV-1 envelope protein gp160,
CC resulting in HIV-1 retention in the ER and thereby preventing viral
CC replication. In a specific embodiment, the chimeric CD4 molecule
CC comprises CD4 fused to 15 amino acids of the C-terminal domain of the T
CC cell receptor CD3epsilon chain; this chimeric CD4 molecule is designated
CC CD4epsilon1015 (ADA44807). A known chimeric CD4 of similar structure but
CC containing only 10 amino acids from CD3epsilon1015 can also be used.
CC Compositions of the invention have an in trans effect on the replication
CC of HIV-1, and may be used to treat and prevent HIV-1 infection. The
CC present sequence represents the chimeric CD4 molecule CD4epsilon1015, which
CC is specifically claimed for use in compositions of the invention.
XX
SQ Sequence 473 AA;
XX
Alignment Scores:
Pred. No.: 232
Score: 8.00 Length: 473
Percent Similarity: 100.00% Matches: 8
Best Local Similarity: 100.00% Conservative: 0
Query Match: 1.20% Mismatches: 0
DB: Indels: 0
Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x ADA44807 (1-473)
OY 1819 GACAGAGCTGACTGCGCGGAGC 1842
XXXXXXXXXXXXXXXXXXXXXXXXXXXX

Db 78 AspaRgAlaSpSerArGArGSer 85

RESULT 30

AAV59170

ID AAV59170 standard; protein; 474 AA.

XX

AC AAV59170;

XX

DT 14-MAR-2000 (first entry)

XX

DE CD4-Ig fusion protein CD4Mmu.

XX

KM HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;

KW secreted protein; SIV infection; medicament.

XX

OS Synthetic.

XX

OS Homo sapiens.

XX

PN CAL340741-C.

XX

PD 14-SEP-1999.

XX

PF 20-JAN-1989; 89CA-00588749.

XX

PR 20-JAN-1989; 89CA-00588749.

XX

PA (GEHO) GEN HOSPITAL CORP.

XX

PI Seed B;

XX

DR WPI; 2000-063015/06.

DR N-PSDB; AAZ48203.

PT New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in

PT the treatment of HIV or simian immunodeficiency virus infections.

XX

PS Example 1; Page 47-53; 89pp; English.

XX

CC The invention provides a fusion gene encoding a fusion protein that

CC comprises an extracellular CD4 DNA sequence or its fragment which binds

CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA

CC sequence of an Ig heavy or light chain, where the DNA sequence encoding

CC the variable region has been replaced with the DNA sequence which encodes

CC extracellular CD4 or its gp120 binding fragment. The fusion protein is

CC capable of being secreted. The fusion proteins are useful for treating

CC HIV or SIV infections in animals, preferably humans. They are also useful

CC for producing medicaments which can be used for treating HIV or SIV

CC infections in humans. The present sequence represents the fusion protein

CC CD4Mmu where the CD4 is linked to human IgG1 at the MscI site upstream of

CC the CH1 region

XX

SQ Sequence 474 AA;

XX

Alignment Scores:

Pred. No.: 232 Length: 474

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.20% Indels: 0

DB: 3 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AAV59170 (1-474)

QY 1819 GACAGAGCTGACTCGCGCGAGC 1842

Db 78 AspaRgAlaSpSerArGArGSer 85

RESULT 31

AAV93011

ID AAV93011 standard; protein; 481 AA.

XX

AC AAV93011;

XX

DT 25-MAR-2003 (revised)

DT 03-AUG-1992 (first entry)

XX

DE Genetic construct which encodes CD4 linked to human IGM at the Pst site

DE upstream of the CH2 region (fusion protein CD4Pmu).

XX

KM Fusion protein; immunoglobulin-like molecule; HIV; SIV; therapy;

KW diagnosis; CD4; gp120; binding fragment; glycoprotein; variable region.

XX

OS Homo sapiens.

XX

PN EP325262-A.

XX

PD 26-JUL-1989.

XX

PF 20-JAN-1989; 89EP-00100913.

XX

PR 22-JAN-1988; 88US-00147351.

XX

PA (GEHO) GEN HOSPITAL CORP.

XX

PI Seed B;

XX

DR WPI; 1989-214472/30.

DR N-PSDB; AAN90359.

PT Immunoglobulin-CD4 fusion proteins - used for treating HIV or SIV

PT infections or detecting HIV or SIV in sample.

XX

PS Example; Table 4, Page 41-47; 68pp; English.

XX

CC The fusion protein genes of the invention pref. comprises cDNA sequences

CC which encode CD4 or a fragment which binds gp120 ligated to an expression

CC plasmid which encodes an antibody in which the variable region of the

CC gene has been deleted (see W087-02671). The CD4 portion of the fusion

CC protein may comprise the complete CD4 sequence, the 370 AA extracellular

CC region and the membrane spanning domain, or the extracellular region. The

CC Ig heavy chain is pref. from IGM, IgG1 or IgG3. The following are

CC specifically claimed: fusion proteins CD4lambda1, CD4Mmu, CD4Pmu,

CC CD4lambda1, and CD4Mmu (No. 67608), PCDA4lambda1 (No. 67609) and

CC PCDA4lambda1 (No. 67610). (Updated on 25-MAR-2003 to correct PA field.)

XX

SQ Sequence 481 AA;

XX

Alignment Scores:

Pred. No.: 232 Length: 481

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.20% Indels: 0

DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AAV93011 (1-481)

QY 1819 GACAGAGCTGACTCGCGCGAGC 1842

Db 78 AspaRgAlaSpSerArGArGSer 85

RESULT 32

AAV19510

ID AAV19510 standard; protein; 481 AA.

XX

AC AAV19510;

XX

DT 09-JAN-2001 (first entry)

XX

DE CD4-IgM fusion protein CD4Pmu.

XX

KM CD4; IGM, human; CD4Pmu; fusion protein; immunoglobulin; HIV; SIV; gp120;

KW therapy; diagnosis.

XX

OS Homo sapiens.

XX

PT Novel neutralizing bispecific fusion proteins effective in viral such as
 PT HIV neutralization, comprises two different binding domains, inducing-
 PS binding domain and induced-binding domain functionally linked by linker.
 XX Claim 39; Page 46-47; 55pp; English.

CC sCD4-SCFv(17b) is a neutralising bispecific fusion protein capable of
 CC binding to two sites of its target protein. The protein comprises a first
 CC binding domain capable of binding to an inducing site on the target
 CC protein, a second binding domain capable of forming neutralising complex
 CC with an induced epitope of the target protein and a linker connecting the
 CC binding domains. sCD4-SCFv(17b) comprises a soluble CD4 fragment
 CC (containing domains D1 and D2) fused to a single chain Fv portion of
 CC antibody 17b via a linker. sCD4-SCFv(17b), its variant, analogue or
 CC mimetic is used for inactivating gp120 protein of HIV, and for
 CC neutralising HIV. It is also used for blocking and preventing the binding
 CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte
 CC CD4 and for inhibiting HIV replication. The chimeric proteins is
 CC particularly useful for treating HIV infection and also AIDS. It is also
 CC after HIV exposure (e.g., mother/infant transmission, post-exposure
 CC prophylaxis, and as a topical inhibitor) and for providing long term
 CC resistance to HIV infections and AIDS. Gene therapy is used to secrete
 CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal
 CC or oral mucosa. The fusion proteins is highly potent, broadly cross-
 CC reactive with neutralising antibody with high in vivo activity and no Fc-
 CC mediated undesirable labelling properties. When the fusion protein is
 CC substantially derived from human proteins, it has minimal immunogenicity
 CC and toxicity in humans which is of great value in prevention of infection
 CC during or immediately after HIV exposure
 XX

SQ Sequence 507 AA;

Alignment Scores:
 Pred. No.: 230 Length: 507
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 3 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AAB00158 (1-507)

OY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
 DB |||||||
 78 AspargalaapSerArgRser 85

RESULT 37

AAW26792
 ID AAW26792 standard; protein; 509 AA.
 XX
 AC AAW26792;

DT 22-JUN-1998 (first entry)
 XX

DE Mouse TIE-2 receptor ligand-3.
 XX

KM TIE-2 receptor ligand; TIE ligand-3; tyrosine kinase; mouse;
 KM angiogenesis; vascularisation; blood vessel growth; tumour; therapy;
 XX receptorbody.
 XX

OS Mus musculus.
 XX

PN WO9748804-A2.
 XX

PD 24-DEC-1997.
 XX

PF 19-JUN-1997; 97WO-US010728.
 XX

PR 19-JUN-1996; 96US-00665926.
 PR 02-JUL-1996; 96US-0021087P.
 PR 02-AUG-1996; 96US-0022999P.
 XX

PT

PA (REG-) REGENERON PHARM INC.
 XX Valenzuela DM, Jones PF, Yancopoulos GD,
 PI
 XX WPI: 1998-063143/06.
 DR N-PSDB; AAT99594.
 XX

PT DNA encoding TIE ligands 3 and 4 - useful for, e.g. blocking blood vessel
 PT growth and promoting neovascularisation, etc.
 XX

PS Example 10; Fig 6A-B; 86pp; English.
 XX

CC This protein comprises novel mouse TIE ligand-3 (mTL-3), a TIE-2 receptor
 CC ligand that is a tyrosine kinase with immunoglobulin and epidermal growth
 CC factor homology domains. Its amino acid sequence was deduced from an
 CC isolated cDNA clone (see AAT99594). A host-vector system for the
 CC production of mTL-3 is claimed, as well as an antibody which specifically
 CC binds mTL-3, a receptorbody which specifically binds mTL-3, a conjugate
 CC comprising mTL-3 and a cytotoxic agent, and an isolated nucleic acid (see
 CC AAT99593) coding for human TIE ligand 4 (see AAW26791). TL3, TL4, their
 CC antagonists, etc. are useful for blocking blood vessel growth, promoting
 CC neovascularisation, promoting or blocking the growth or differentiation
 CC of a cell expressing the TIE receptor and attenuating or preventing
 CC tumour growth in humans
 XX

SQ Sequence 509 AA;

Alignment Scores:
 Pred. No.: 230 Length: 509
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AAW26792 (1-509)

OY 1083 CGTGCCGAGCGTCCCGCGCGCA 1106
 DB |||||||
 86 ArgalaatlnArglaatlnArgla 93

RESULT 38

AAW47533
 ID AAW47533 standard; protein; 509 AA.
 XX
 AC AAW47533;

DT 09-SEP-1998 (first entry)
 XX

DE Amino acid sequence of murine TIE ligand 3.
 XX

KM Chimeric TIE ligand; TIE-2 ligand; neovascularisation; tumour; mouse.
 KM Mus sp.
 XX

PN WO9805779-A1.
 XX

PD 12-FEB-1998.
 XX

PF 01-AUG-1997; 97WO-US013557.
 XX

PR 02-AUG-1996; 96US-0022999P.
 PR 25-OCT-1996; 96US-00740223.
 XX

PA (REG-) REGENERON PHARM INC.
 XX

PI Davis S, Yancopoulos GD,
 XX

XX WPI: 1998-145615/13.
 DR N-PSDB; AAV18620.
 XX

PT Modified human TIE-2 receptor ligand(s) - useful for promoting wound
 PT healing.

XX Example 20; Fig 21; 202pp; English.
XX
CC This is the amino acid sequence of the murine TIE ligand 3, used in the
CC method of the invention, involving the production of TIE-2 ligands which
CC promote healing. The nucleic acids, vectors and host cells used in the
CC method of the invention are useful for the recombinant production of the
CC ligands. The ligands, etc. are useful for blocking blood vessel growth,
CC promoting neovascularisation, promoting the growth or differentiation of
CC a cell expressing the TIE receptor, blocking the growth or
CC differentiation of a cell expressing the TIE receptor and for attenuating
CC or preventing tumour growth in a human
XX
SQ Sequence 509 AA;
XX
Alignment Scores:
Pred. No.: 230 Length: 509
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: Gaps: 0
XX
US-10-029-345A-108_COPY_532_2532 (1-2001) x AAM47533 (1-509)
QY 1083 CGTGGCCAGCGTGCAGCGTGCA 1106
DB 86 ArgAlaGlnArgAlaGlnArgAla 93
XX
RESULT 39
ID AAY23734 standard; protein; 509 AA.
XX
AC AAY23734;
XX
DT 08-SEP-1999 (first entry)
XX
DE TIE ligand-3 (TL3) amino acid sequence.
XX
KM TIE ligand-3; TL3; angiogenic factor; receptor tyrosine kinase; AR-1;
KM angiogenesis regulator; neoplastic disease; tumour angiogenesis;
KM wound healing; thromboembolic disease; atherosclerosis;
KM inflammatory disease; AR-1 receptor.
XX
OS Mus sp.
XX
PN WO9932639-A1.
XX
PD 01-JUL-1999.
XX
PF 17-DEC-1998; 98WO-US026800.
XX
PR 19-DEC-1997; 97US-00994309.
XX
PA (REGG-) REGENERON PHARM INC.
XX
PI Valenzuela DM;
XX
XX WPI; 1999-418933/35.
DR N-PSDB; AAX85781.
XX
PT Receptor tyrosine kinase, AR-1 and related nucleic acid molecules.
XX
XX Example 1; Fig 1A-C; 50pp; English.
XX
CC The present sequence represents TIE ligand-3 (TL3), an angiogenic factor.
CC The specification describes a receptor tyrosine kinase, designated AR-1,
CC which is related to TIE ligands. AR-1 is believed to be a regulator of
CC angiogenesis and thus the factor, as well as nucleic acids encoding it,
CC are useful in the diagnosis and treatment of certain diseases such as
CC neoplastic diseases involving tumour angiogenesis, wound healing,
CC thromboembolic diseases, atherosclerosis and inflammatory diseases. AR-1
CC can also be used to support the survival and/or growth and/or migration

CC and/or differentiation of human AR-1 receptor expressing cells. AR-1 can
CC be used to identify the AR-1 receptor. AR-1-cytotoxic conjugates can be
CC used to target tumours expressing AR-1 or its receptor. Antagonists of AR
CC -1, e.g. antibodies are useful to prevent or attenuate its biological
CC activity
XX
SQ Sequence 509 AA;
XX
Alignment Scores:
Pred. No.: 230 Length: 509
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: Gaps: 0
XX
US-10-029-345A-108_COPY_532_2532 (1-2001) x AAY23734 (1-509)
QY 1083 CGTGGCCAGCGTGCAGCGTGCA 1106
DB 86 ArgAlaGlnArgAlaGlnArgAla 93
XX
RESULT 40
ID AAY90398
XX
AC AAY90398; standard; protein; 509 AA.
XX
DT 12-FEB-2001 (first entry)
XX
DE Mouse TIE ligand-3 protein sequence.
XX
KM AR-2 fibrinogen-like domain; angiotensin related-2; mouse; ischemia;
KM diabetes; tumour angiogenesis; neoplastic disease; atherosclerosis;
KM thromboembolic disease; inflammatory disease; wound healing;
KM vascularisation; therapy; diagnosis; TIE ligand-3.
XX
OS Mus sp.
XX
PN WO200052167-A1.
XX
PD 08-SEP-2000.
XX
PF 10-FEB-2000; 2000WO-US003381.
XX
PR 02-MAR-1999; 99US-0122499P.
XX
PA (REGG-) REGENERON PHARM INC.
XX
PI Jones PF, Valenzuela DM;
XX
XX WPI; 2000-579286/54.
DR N-PSDB; AAA37838.
XX
PT Novel nucleic acid molecules encoding fibrinogen-like domain of
PT Angiotensin Related-2 factor useful for treating ischemia, diabetes, and
PT for wound healing.
XX
XX Example 1; Fig 1; 55pp; English.
XX
CC This sequence is the mouse TIE ligand-3. This sequence was used in the
CC isolation of the angiotensin related-2 (AR-2) fibrinogen-like domain
CC (FD) of the invention. AR-2 is useful for treating and diagnosing
CC ischemia, diabetes, tumour angiogenesis, neoplastic diseases,
CC thromboembolic diseases, atherosclerosis, inflammatory diseases, and for
CC wound healing. The DNA is useful for developing ligands, screening
CC agonists and antagonists of AR-2, and as a therapeutic for treating
CC disorders involving cells, tissues or organs expressing AR-2 receptor. AR
CC -2 is useful to promote the growth, survival, migration, stabilisation or
CC destabilisation, and/or differentiation of cells expressing AR-2
CC receptor. AR-2 is also useful in assay systems to identify agonists and
CC antagonists of AR-2 receptor. AR-2 is also useful for inducing or
CC preventing vascularisation in diseases or disorders where such function

CC is indicated, for delivering toxins to a receptor bearing cells and as
CC diagnostic reagents for detecting the disease by tissue staining or whole
CC body imaging
XX
SQ Sequence 509 AA;

Alignment Scores:

Pred. No.:	230	Length:	509
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	3	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AAY90398 (1-509)

QY 1083 CGTCCAGCGTCCAGCGTGCA 1106

Db 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 41

ADD69365

ID ADD69365 standard; protein; 509 AA.

AC ADD69365;

XX 15-JAN-2004 (first entry)

DE Human Ang-3 protein - SEQ ID 101.

XX fusion protein; fibrinogen-like; coiled-like domain;

KW angiotensin-related factor; ARF; angiotensin; Ang-1; Ang-2; Ang-3;

KW Ang-4; Ang-2X; vulnerable; antiinflammatory; vasotropic; necrosis;

KW ischaemia; inflammation; wound healing; CCD; FLD; human.

XX Homo sapiens.

PN WO2003048185-A2.

XX 12-JUN-2003.

PF 21-NOV-2002; 2002WO-US037660.

PR 30-NOV-2001; 2001US-0334488P.

XX (GENV-) GENVEC INC.

PA Kessler PD, Kovacs I;

XX WPI; 2003-513736/48.

DR WPI; 2003-513736/48.

XX New fusion protein comprising a fibrinogen-like or coiled-like domain,

PT useful for preparing a composition for treating necrosis, ischemia or

PT inflammation, or for promoting wound healing.

PS Disclosure; SEQ ID NO 101; 340pp; English.

XX The invention relates to a novel fusion protein comprising a fibrinogen-

CC like domain (FLD) or coiled-coil domain (CCD). The domain may be

CC identical or homologous to that of an angiotensin-related factor (ARF),

CC examples of which include Ang (angiotensin)-1, Ang-2, Ang-3, Ang-4 and

CC Ang-2X. The molecules of the invention demonstrate vulnerable,

CC antiinflammatory and vasotropic activities whilst the fusion protein may

CC be useful for preparing a composition for treating necrosis, ischemia or

CC inflammation, as well as for promoting wound healing. The current

CC sequence is that of the human Ang-3 protein of the invention.

XX Sequence 509 AA;

Alignment Scores:

Pred. No.:	230	Length:	509
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	7	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x ADD69365 (1-509)

QY 1083 CGTCCAGCGTCCAGCGTGCA 1106

Db 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 42

AAR20152

ID AAR20152 standard; protein; 519 AA.

XX AAR20152;

XX 25-MAR-2003 (revised)

DT 31-MAR-1992 (first entry)

XX Human CD4 sequence encoded by PATY.6.

DE Human immunodeficiency virus; HIV; gp 120; AIDS; ARC; glycoprotein,

KW acquired immune deficiency syndrome; AIDS related complex;

KW T helper lymphocytes.

XX Homo sapiens.

OS Homo sapiens.

XX Key

FH Location/Qualifiers

FT Peptide 1..25

XX /label= signal_sequence

PN WO9118618-A.

XX 12-DEC-1991.

PD 25-MAY-1990; 90US-00529186.

PF 25-MAY-1990; 90US-00529186.

PR 25-MAY-1990; 90US-00529186.

XX (BIOJ) BIOGEN INC.

PA Fisher RA, Hession C, Burkly LC;

XX WPI; 1992-007200/01.

DR N-PSDB; AAQ20327.

XX New immuno-therapeutic human CD4 variants and derivs. - elicit AB

PT production to HIV gp.120, useful in treating, preventing and diagnosing

PT AIDS, ARC and HIV infections.

XX Disclosure; Fig 28; 179pp; English.

PS The sequence was deduced from the DNA sequence of subclone PATY.6, contg.

CC DNA coding for the full-length human CD4. The clone was constructed from

CC plasmids PB8178A and PB8378 (both in US8802940). The DNA can be used to

CC express recombinant CD4 and analogues for use in diagnosis and treatment

CC of diseases caused by infective agents whose primary targets are T4+

CC lymphocytes. See also AAR20148-R20155 and AAR21078. (Updated on 25-MAR-

CC 2003 to correct PA field.)

XX Sequence 519 AA;

Alignment Scores:

Pred. No.:	229	Length:	519
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	2	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AAR20152 (1-519)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842

Db 139 AspaRgAlaSpSerArgSer 146

|||||

RESULT 43

AAP94703

ID AAP94703 standard; protein; 524 AA.

XX

XX AAP94703;

AC

XX 25-MAR-2003 (revised)

DT 22-MAR-1991 (first entry)

XX

DE Sequence encoded by T4 lymphocyte cDNA obtained from PBL clone lambda-203

DE -4.

XX

XX HIV; soluble T4; immunotherapeutic; prophylactic; diagnostic; AIDS; ARC.

XX

XX Homo sapiens.

XX

PH Key Location/Qualifiers

FT Misc-difference 67 /note= "AA DESIGNATED NUMBER -23"

FT Misc-difference 90 /note= "AA DESIGNATED NUMBER 1"

FT Misc-difference 92 /note= "MATURE N-TERMINUS"

FT

PN WO8901940-A.

XX

PD 09-MAR-1989.

XX

PF 01-SEP-1988; 88WO-US002940.

XX

XX 04-SEP-1987; 87US-00094322.

PR 07-JAN-1988; 88US-00141649.

XX

PA (BIOJ) BIOGEN INC.

XX

PI Fisher RA, Gilbert W, Sato VL, Flavell RA, Maraganore JM;

DR MPI; 1989-085519/11.

DR N-PSDB; AAN90642.

XX

PT DNA sequences coding for soluble T4-like polypeptide(s) - used in

PT immuno:therapeutic and immunosuppressive compns. and for preventing,

PT treating or detecting AIDS.

XX

XX

PS Disclosure; Fig 3; 207pp; English.

XX

CC The polypeptides encoded are useful in immunotherapeutic, prophylactic

CC and diagnostic compns. They can be used to purify HIV from a sample. The

CC soluble T4 protein-based compns. are useful in treating immunodeficient

CC patients suffering from diseases caused by agents whose primary targets

CC are T4+ lymphocytes. They can be used for preventing, treating or

CC detecting AIDS, ARC and HIV infection. (Updated on 25-MAR-2003 to correct

CC PR field.)

XX

XX SQ Sequence 524 AA;

Alignment Scores:

Pred. No.: 229 Length: 524

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.20% Indels: 0

DB: 1 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x AAP94703 (1-524)

Qy 1819 GACAGAGCTGACTGCGCGCGAGC 1842

Db 144 AspaRgAlaSpSerArgSer 151

RESULT 44

AAG51289

ID AAG51289 standard; protein; 525 AA.

XX

XX AAG51289;

AC

XX 18-OCT-2000 (first entry)

DT

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 65080.

XX

XX Protein identification; signal transduction pathway; metabolic pathway;

XX hybridization assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

XX 25-FEB-2000; 2000EP-00301439.

PF

XX 25-FEB-1999; 99US-0121825P.

XX

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126785P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128334P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143622P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.

PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153707P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0155659P.
PR 29-SEP-1999; 99US-0156458P.
PR 04-OCT-1999; 99US-0156596P.
PR 05-OCT-1999; 99US-0157117P.
PR 06-OCT-1999; 99US-0157553P.
PR 07-OCT-1999; 99US-0157865P.
PR 08-OCT-1999; 99US-0158029P.
PR 12-OCT-1999; 99US-0158232P.
PR 13-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	229	Length:	525
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	3	Gaps:	0

US-10-029-345a-108_COPY_532_2532 (1-2001) x AAG51289 (1-525)

QY 374 GTGGTTTCTAGTTCTCTCGTT 397

Db 508 ValGlyLeuLeuSerSerLeuVal 515

RESULT 45
AAG22387

[illegible]

Db	78	AsparGalaAsparAIGarSer	85
RESULT 47			
ID	AA046679	standard; protein; 530 AA.	
XX	AA046679;		
AC			
XX			
DT	25-MAR-2003	(revised)	
DT	08-AUG-1994	(first entry)	
XX			
DE	CD4-IgG2 chimeric heavy chain.		
XX			
KM	CD4; gamma; heavy chain; chimeric; chimaeric; immunoconjugate; HIV;		
KM	human immunodeficiency virus; radionuclide; toxin; therapy; treatment;		
XX	imaging; detection; targeting; immunoglobulin; IgG.		
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	1. .204		
FT	/label= CD4 Region.		
FT	205. .302		
FT	/label= CH1 Region.		
FT	303. .314		
FT	/label= Hinge Region.		
FT	315. .423		
FT	/label= CH2 Region.		
FT	424. .530		
FT	/label= CH3 Region.		
XX			
PN	WC9403191-A1.		
PD	17-FEB-1994.		
XX			
PP	06-AUG-1993; 93WC-US007422.		
XX			
PR	07-AUG-1992; 92US-00927931.		
XX			
PA	(PROG-) PROGENICS PHARM INC.		
PI	Allaway GP, Maddon PJ;		
XX			
DR	WPI; 1994-065392/08.		
DR	N-PSDB; AA055751.		
XX			
PT	Non-peptidyl toxin or radionuclide and CD4-gamma 2 or CD4-IgG2		
PT	Immunoconjugates - used to kill HIV-infected cells and to image and		
PT	stage HIV infection.		
PS	Disclosure; Fig 4; 142pp; English.		
XX			
CC	A tetramer comprising 2 IgG2 heavy chains or two CD4-IgG2 chimeric heavy		
CC	chains and two kappa light chains or CD4-kappa light chains (AA046680)		
CC	linked to a non-peptidyl toxin or a gamma radiation-emitting radionuclide		
CC	of low to moderate cytotoxicity. The resulting immunoconjugate comprising		
CC	the toxin can be used to kill HIV infected cells and to treat HIV		
CC	infected subjects to reduce the population of HIV infected cells. It can		
CC	also be used to reduce the likelihood of infection. The immunoconjugate		
CC	comprising the radionuclide can be used to image HIV infected tissue, to		
CC	calculate the stage of HIV infection or the efficacy of an anti-HIV		
CC	treatment using the imaging technique and for determining the prognosis		
CC	of an HIV infected subject. (Updated on 25-MAR-2003 to correct PN field.)		
XX			
XX	Sequence 530 AA;		
Alignment Scores:			
Pred. No.:	229	Length:	530
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0

	DB:		2	Gaps:	0
US-10-029-345A-108_COPY_532_2532 (1-2001) x AAR46679 (1-530)					
OY	1819	GACAGAGCTGACTCGCGCGAGC	1842		
Dd	78	AapAtgAlaaPSeRrArgRser	85		
RESULT 48					
ID	AAy85080	standard; protein; 530 AA.			
XX	AAy85080;				
AC	AAy85080;				
XX					
DT	19-JUN-2000	(first entry)			
XX					
DE	CD4-IgG2 chimeric heterotetramer heavy chain amino acid sequence.				
KM	CD4-IgG2 chimeric heavy chain heterotetramer; immunoonjugate; treatment;				
KW	cytotoxic radionuclide; cell surface glycoprotein; prevent; infection;				
KM	cellular immune response interaction mediator; HIV interaction; staging;				
KW	prognosis; envelope glycoprotein burden; human.				
OS	Homo sapiens.				
XX					
PN	US6034223-A.				
PD	07-MAR-2000.				
XX					
PE	07-JUN-1995;	95US-00477460.			
XX					
PR	07-AUG-1992;	92US-00927931.			
PR	06-AUG-1993;	93WO-US007442.			
XX	03-FEB-1995;	95US-00379516.			
PA	(PROG-) PROGENICS PHARM INC.				
PJ	Allaway GP, Maddon PJ;				
DR	WPI: 2000-269502/23.				
XX	N-PESDB; AAZ98856.				
PT	New immunoconjugate, used to treat, prevent or image human immune				
PT	deficiency virus infection, comprises radionuclide attached to				
XX	heterotetramer of CD4-immunoglobulin chimeras.				
PS	Disclosure; Fig 4; 58pp; English.				
XX					
XX	This sequence represents the CD4-IgG2 chimeric heavy chain amino acid				
CC	sequence from the CD4-IgG2 chimeric heterotetramer. The invention relates				
CC	to an immunoconjugate comprising a cytotoxic radionuclide and a				
CC	heterotetramer of two heavy chains and two light chains. The cytotoxic				
CC	radionuclide is linked to either the heavy chains or the light chains, on				
CC	to all four chains, directly or through a bifunctional chelator. Both				
CC	heavy chains are chimeric CD4-Ig (immunoglobulin) G2 chains encoded by				
CC	vector CD4-IgG2HC-pRCMV (ATCC 75193) and both light chains are chimeric				
CC	CD4-kappa chains encoded by vector CD4-kLC-pRCMV (ATCC 75194). CD4 is a				
CC	non-polymorphic cell surface glycoprotein that is expressed on the				
CC	surface of helper T lymphocytes, cells of the monocyte/macrophage lineage				
CC	and dendritic cells. CD4 associates with major histocompatibility complex				
CC	(MHC) class II molecules on the surface of antigen presenting cells to				
CC	mediate efficient cellular immune response interactions. In humans CD4 is				
CC	the target of interaction with the human immunodeficiency virus HIV. The				
CC	immunoconjugate is used to kill cells infected with HIV, and for treating				
CC	or preventing infection. It is also used for imaging HIV-infected tissues				
CC	(for staging or prognosis of infection, and for assessing efficacy of				
CC	treatments). The immunoconjugate is also used to determine the HIV				
CC	envelope glycoprotein burden, once determined, this information is used				
CC	in the staging and prognosis of HIV infected patients. The				
CC	immunoconjugate should be active against all strains of HIV (since the				
CC	CD4-gp120 interaction is essential for infection). The heterotetramers				
CC	are assembled intracellularly and secreted efficiently from mammalian				
CC	cells, allowing high recovery and purification from the culture medium.				

CC They have longer half-life in serum and greater avidity than heavy chain
 CC dimers
 XX
 SQ Sequence 530 AA;

Alignment Scores:

Pred. No.: 229 Length: 530
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 3 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AAB85080 (1-530)

QY 1819 GACAGAGCTGACTCGGCGGAGC 1842

DB 78 ASPARGALASPSETRARGTSER 85

RESULT 49

AAB67323 ID AAB67323 standard; protein; 530 AA.

AC AAB67323;

XX 23-APR-2001 (first entry)

DE CD4-IgG2 chimeric heavy chain protein.

XX Immunocjugate; chelator; chimeric; HIV; human immunodeficiency virus.

OS Homo sapiens.

XX US6177549-B1.

XX 23-JAN-2001.

XX 10-JUN-1999; 99US-00329916.

XX 07-AUG-1992; 92US-00927931.

XX 06-AUG-1993; 93WO-US007422.

XX 03-FEB-1995; 95US-00379516.

XX 07-JUN-1995; 95US-00477460.

XX (PROG-) PROGENICS PHARM INC.

XX Maddon PJ, Allaway GP;

XX WPI; 2001-158582/16.

XX Immunocjugate for treating human immunodeficiency virus-infected

XX subject, consists of cytotoxic radionuclide linked to heterotetramer

XX comprising two chimeric CD4-IgG2 heavy chains and two chimeric CD4-

XX kappa light chains.

XX Disclosure; Fig 4; 43pp; English.

XX The present invention relates to an immunocjugate, comprising a

XX cytotoxic radionuclide linked, directly or via a bifunctional chelator,

XX to a heterotetramer having two chimeric CD4-IgG2 heavy chains encoded by

XX an expression vector CD4-IgG2HC-PRCMV and two chimeric CD4-kappa light

XX chains encoded by an expression vector CD4-kLC-PRCMV. The invention is

XX useful for killing human immunodeficiency virus (HIV)-infected cells, for

XX the treatment and prevention of infection with HIV

XX Sequence 530 AA;

Alignment Scores:

Pred. No.: 229 Length: 530
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0

DB: 4 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AAB67323 (1-530)

QY 1819 GACAGAGCTGACTCGGCGGAGC 1842

DB 78 ASPARGALASPSETRARGTSER 85

RESULT 50

AAB80884 ID AAB80884 standard; protein; 530 AA.

XX AAB80884;

XX 29-MAY-2001 (first entry)

DE Human CD4-IgG2 chimeric heavy chain.

XX Human; Anti-HIV; CD4-IgG2 chimeric heterotetramer;

XX immunoglobulin gamma 2.

XX Homo sapiens.

XX US6187748-B1.

XX 13-FEB-2001.

XX 07-JUN-1995; 95US-00485372.

XX 08-FEB-1991; 91US-00653684.

XX 10-FEB-1992; 92WO-US001143.

XX 08-DEC-1992; 92US-00960440.

XX (PROG-) PROGENICS PHARM INC.

XX Maddon PJ, Beaudry GA;

XX WPI; 2001-264981/27.

XX N-PSDB; AAF77830.

XX Inhibiting human immunodeficiency virus (HIV) infection of a CD4+ cell,

XX or treating a subject having CD4+ cells infected with HIV involves using

XX CD4-IgG2 chimeric heterotetramer to form a complex with the HIV.

XX Disclosure; Fig 4; 55pp; English.

XX The present invention relates to a method for inhibiting infection of a

XX CD4+ cell by HIV. The method comprises contacting the HIV with a CD4-IgG2

XX chimeric heterotetramer to form a complex with the HIV (CD4 = cluster of

XX differentiation 4; IgG2 = immunoglobulin gamma 2). CD4 is a cell surface

XX glycoprotein that is expressed primarily on the surface of T cells. In

XX two heavy and two light chains which are encoded by expression vectors

XX CD4-IgG2HC-PRCMV (V1) and CD4-kLC-PRCMV (V2), respectively. The method

XX is used to inhibit infection of a CD4+ cell by a HIV and to prevent CD4+

XX cells of a subject from becoming infected with HIV. The method is also

XX useful for treating a subject having CD4+ cells infected with HIV. The

XX present sequence is human fusion protein: CD4-Ig2 chimeric heavy chain of

XX the CD4-IgG2 chimeric heterotetramer. This sequence was used in the

XX method of the present invention

SQ Sequence 530 AA;

Alignment Scores:

Pred. No.: 229 Length: 530
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 4 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x AAB80884 (1-530)

Qy 1819 GACAGACTGACTCGCGCGGAGC 1842
|||
Db 78 AspArgAlaAspSerArgSer 85

Search completed: June 21, 2004, 13:09:22
Job time : 169.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: June 21, 2004, 13:09:27 ; Search time 93.5 Seconds

(without alignments)
12083.628 Million cell updates/sec

Title: US-10-029-345A-108_COPY_532_2532

Perfect score: 667
Sequence: 1 gttgcatcgtgcccacatgagat.....tggaaatcatctgagctcc 2001

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1163542 seqs, 282313646 residues

Word size: 1

Total number of hits satisfying chosen parameters: 349658

Minimum DB seq length: 473

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODEL=frame+npz.model -DEV=xld
-Q=/cgn2_1/USPTO.spool.p/US10029345/runat.21062004.124617.5758/app.query.fasta_1.2183
-DB=Published Applications AA -QFMT=fastaan -SUFFIX=oligo.rapb -MINMATCH=0.1
-LOOPCT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=oligo
-TRANS-human40.cdt -LIST=75 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=50 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=473
-MAXLEN=200000000 -USER=US10029345 @CGN_1.1.13 @runat.21062004.124617.5758
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

1	472	70.8	517	9	US-09-964-277-21	Sequence 21, Appl
2	472	70.8	665	9	US-09-816-494-2	Sequence 2, Appl
3	472	70.8	665	9	US-09-964-277-2	Sequence 2, Appl
4	472	70.8	665	12	US-10-072-012-680	Sequence 680, App
5	472	70.8	665	12	US-10-072-012-681	Sequence 681, App
6	472	70.8	665	12	US-10-168-506-14	Sequence 14, Appl
7	472	70.8	665	12	US-10-343-357-7	Sequence 7, Appl
8	472	70.8	665	15	US-10-094-749-2312	Sequence 2312, Ap
9	472	70.8	665	15	US-10-377-077-26	Sequence 26, Appl
10	472	70.8	665	16	US-10-257-028-2	Sequence 2, Appl
11	472	70.8	665	16	US-10-648-593-240	Sequence 240, App
12	472	70.8	665	16	US-10-648-593-247	Sequence 247, App
13	472	70.8	665	12	US-10-072-012-679	Sequence 679, App
14	472	70.8	665	12	US-10-072-012-703	Sequence 703, App
15	472	70.8	665	12	US-10-425-114-54204	Sequence 54204, A
16	400	60.0	662	12	US-10-072-012-258	Sequence 258, App
17	394	59.1	660	12	US-10-072-012-256	Sequence 256, App
18	258	38.7	660	12	US-10-296-115-1259	Sequence 1259, Ap
19	67	10.0	660	12	US-10-072-012-682	Sequence 682, App
20	64	9.6	677	12	US-10-072-012-683	Sequence 683, App
21	20	3.0	501	12	US-10-072-012-702	Sequence 702, App
22	20	3.0	625	12	US-10-072-012-659	Sequence 699, App
23	20	3.0	663	12	US-10-072-012-700	Sequence 700, App
24	12	1.6	616	12	US-10-072-012-266	Sequence 266, App
25	10	1.2	498	12	US-10-282-122A-57268	Sequence 57268, A
26	8	1.2	503	14	US-10-225-060-11	Sequence 11, Appl
27	8	1.2	509	12	US-10-214-812-2	Sequence 2, Appl
28	8	1.2	509	12	US-10-214-812-3	Sequence 3, Appl
29	8	1.2	509	14	US-10-215-224-2	Sequence 2, Appl
30	8	1.2	509	14	US-10-215-224-3	Sequence 3, Appl
31	8	1.2	509	14	US-10-225-060-10	Sequence 10, Appl
32	8	1.2	513	9	US-09-833-745-57	Sequence 57, Appl
33	8	1.2	529	12	US-10-424-599-231545	Sequence 231545,
34	8	1.2	530	8	US-08-485-163-5	Sequence 5, Appl
35	8	1.2	530	9	US-09-766-995-4	Sequence 4, Appl
36	8	1.2	532	10	US-09-939-537-6	Sequence 6, Appl
37	8	1.2	532	11	US-09-243-008-6	Sequence 11, Appl
38	8	1.2	544	14	US-10-046-232-22	Sequence 22, Appl
39	8	1.2	544	14	US-10-156-761-11387	Sequence 11387, A
40	8	1.2	555	16	US-10-628-088-415	Sequence 415, Appl
41	8	1.2	575	10	US-09-939-537-4	Sequence 4, Appl
42	8	1.2	575	11	US-09-243-008-4	Sequence 4, Appl
43	8	1.2	590	9	US-09-934-060A-13	Sequence 13, Appl
44	8	1.2	653	14	US-10-062-937B-5	Sequence 5, Appl
45	8	1.2	653	15	US-10-391-399-73	Sequence 73, Appl
46	8	1.2	677	9	US-09-833-745-39	Sequence 39, Appl
47	8	1.2	677	15	US-10-369-493-6901	Sequence 6901, Ap
48	8	1.2	711	9	US-09-792-630-33	Sequence 33, Appl
49	8	1.2	711	10	US-09-953-351-33	Sequence 33, Appl
50	8	1.2	711	13	US-10-080-376-33	Sequence 33, Appl
51	8	1.2	711	14	US-10-082-671-33	Sequence 33, Appl
52	8	1.2	711	14	US-10-097-100-33	Sequence 33, Appl
53	8	1.2	711	14	US-10-023-208-33	Sequence 23, Appl
54	8	1.2	720	9	US-09-934-060A-2	Sequence 2, Appl
55	8	1.2	720	9	US-09-934-060A-4	Sequence 4, Appl
56	8	1.2	774	12	US-10-424-599-257796	Sequence 257796,
57	8	1.2	788	14	US-10-073-118-26	Sequence 26, Appl
58	8	1.2	912	9	US-09-987-482-2	Sequence 2, Appl
59	8	1.2	1222	12	US-10-282-122A-69799	Sequence 69799, A
60	8	1.2	1404	12	US-10-425-114-55163	Sequence 55163, A
61	8	1.2	2630	15	US-10-369-493-3693	Sequence 3693, Ap
62	8	1.2	2883	8	US-08-681-219-32	Sequence 32, Appl
63	8	1.2	2883	9	US-09-987-482-1	Sequence 1, Appl
64	8	1.2	2883	11	US-09-230-111C-30	Sequence 30, Appl
65	8	1.2	2883	12	US-10-392-113-21	Sequence 21, Appl
66	8	1.2	2883	14	US-10-092-133-30	Sequence 30, Appl
67	8	1.2	2883	16	US-10-408-765A-1970	Sequence 1970, Ap
68	8	1.2	2884	12	US-10-267-502-370	Sequence 370, App
69	8	1.2	2884	12	US-10-267-502-372	Sequence 372, App
70	8	1.2	3972	14	US-10-156-761-8476	Sequence 8476, Ap
71	8	1.2	3972	16	US-10-204-862A-4	Sequence 4, Appl
72	8	1.2	3972	16	US-10-204-862A-8	Sequence 8, Appl
73	8	1.2	5532	14	US-10-156-761-8480	Sequence 8480, Ap

c 74 8 1.2 5532 16 US-10-204-862A-7
c 75 7 1.1 473 12 US-10-282-122A-52396

ALIGNMENTS

RESULT 1

US-09-964-277-21
; Sequence 21, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-964-277-21

Alignment Scores:

Pred. No.: 0 Length: 517
Score: 472.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.76% Indels: 0
DB: 9 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-964-277-21 (1-517)

QY 586 ACCGTGCAAGCTGATCTTATCCCGAGTCTCATTTCTGCGTGGCTTGGAATGAC 645
Db 46 ThyrserProlyspProaspPheileProgluserHisPheleuAryValProvalaenasp 65
QY 646 AGCTTTGTGAGAAATTTTGGCGTGTGAGCAAACTCAGTAACTTCAATTGAGAAAGA 705
Db 66 SerPhecygslulysileuProthPheleuAspIysSerValaspPheilegIulysala 85
QY 706 AAAGCTCCCATGAGATGTGTTTCTAGTGCATGTTTACTGGAGTCTCCGCTCCGACAC 765
Db 86 LysalSerasnngIyCysValIleuValHisCysleuValaglYIleSerArgSerAlaThr 105
QY 766 ATGCTATGCTCATGATCATGAGAGAGAGAGAGAGATGTTTATGATGAACTTACAGATT 825
Db 106 ILeAlaIleAlaIlyrIleleuetylsArgmetAspMetSerIeuaSpIuAlaTyArgPhe 125
QY 826 GTGAAAGAAAAAGACCTATATATCTCCAACTTCAATTTTCTGGGCAACTCCGCGAC 885
Db 126 VallygIulysArgProthrIleSerProasnPhasnPheleuGlyGlnIleuueasp 145
QY 886 TATGAAAGAAAGATTGAAGACCAAGATGAGCATCGGSCCAAGAGCAAACTCAAGCTG 945
Db 146 TyGluIulysIlelylsasnngInthrgIyAlaSerGlyProIysSerIysIeulysIeu 165
QY 946 CTGCACTGTGAGAGGCAATGAACTGTCCGCTGTCTGAGAGGTGGAAGAAAAAGC 1005
Db 166 IeuhIleleuGlnIulysProasnGlnProvalProAlaValSerGlnGlyGlnIulysSer 185
QY 1006 GAGAGCCCCCTCATGTCACCTGTGTGCCACTGTGCTACTCTCAGAGGAGAGCAAAAGG 1065
Db 186 GluThrProIeuserProProcyAlaAspSerAlaThrSerGlnAlaIagIyGlnArg 205
QY 1066 CCGCGTCATCCCGGCGAGCTGCCAGCGTGCACAGCGTGCAGCGCTGCTGTAGAGAGC 1125
Db 206 ProvalHisProIaSerValProSerValProSerValGlnProSerIeuleuGlnIulsp 225
QY 1126 AGCCCGCTGTACAGGCGCTCATGTCGAGCTGTCCGAGAGAGGCTGGAAGACAGC 1185

Db 226 SerProIeuaValGlnAlaIeuserGlyIleuHisIleuSerAlaAspArgIeulaspSer 245
QY 1186 AATAAGCTCAAGCGTTCCTTCTCTGATATCAAAATGATTTCATTAATCAGCGACAG 1245
Db 246 AsnIulysIulysArgSerPheSerIeuaSpIleIysSerValSerIySerAlaSerIec 265
QY 1246 GCAGCATCTCTTACATGGCTTCTTCATCATGAAAGATGCTTTGAAATACTACAACCTTCC 1305
Db 266 AlaAlaSerIeuhIagIyPheSerSerSerGlnuAspAlaIeulGlyTytyrIySProSer 285
QY 1306 ACTACTCTGATGGAGCAACAAGCTATGSCAGTTCTCCCGTTCAGAGAACTATCCGAG 1365
Db 286 ThrThrIeuaSpIyThrasnIulysIeucyGlnPheSerProvalGlnIleuSerGlu 305
QY 1366 CAGACTCCGGAACCAAGTCTGTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1425
Db 306 GlnThrProGluThrSerProaspIySerGlnIulAlaSerIleProIulysIeulnThr 325
QY 1426 GCAGGCTTCAGACAGCAAGAGAGAGAGATTTGCATTGGTGCAGAACAGCAGTGGC 1485
Db 326 AlaArgProSerAspSerGlnSerIySarIeuhHisSerValArgThrSerSerGly 345
QY 1486 ACCGCGCAGAGGCTCCCTTTATCTCCAGTCGATGAGAGGAGGAGGAGAGCAATTAC 1545
Db 346 ThrAlaGlnArgSerIeuleuSerProIeuhIasArgSerGlySerValGlnAspAsnTy 365
QY 1546 CACACCAAGCTTCCTTTGGCGCTTTCCACCAAGCAGACAGCACTTCACGAGTCTGGC 1605
Db 366 HisThrSerPheIeuhPheGlyIeuserThrsErGlnIHisIeuthrIySerAlaGly 385
QY 1606 CTGGGCTTAAAGGGGTGGCAGCTCGGATATCTTGGCCCCCGGAGACCTTACCCCTTCC 1665
Db 386 IeugIyIeulysGlyTyPheIleSerAspIleuValaProGlnThrSerThrProSerIeu 405
QY 1666 ACCAGCACTGTATTTTGGCCAGAGAGTCTCAACTTCTACTGCTGCTCAGCCATCTAC 1725
Db 406 ThrsSerIerTyPyrPheAlaThrGluSerSerHisPheTyrsSerAlaSerAlaIleTy 425
QY 1726 GAGGCAATGTCAGTACTCTGCTTCAAGCTGACAGCAGCAGCTGCCACTTGGAGACAA 1785
Db 426 GlyIySerAlaSerIyrsErAlaTySerCyssErGlnuProthrcyGlyAspGln 445
QY 1786 GTCTATCTGTGCCAGGCGGAGAGAGCAAGTGAACAGAGCTGACTGCGCGGAGACTGG 1845
Db 446 ValTyrsErValaIyGlyArgGlnIulysProSerAspArgAlaAspSerArgArgSerTyr 465
QY 1846 CATGAAGAGACCCCTTTGAAGACAGTTTAAACGAGAGAGCTGCCAAATGAAATTTGGA 1905
Db 466 HisGluIulysErProPheGlnIulysGlnPheIyArgArgSerCyGlnMetGluPheGly 485
QY 1906 GAGAGCATTCATGTCAAGAAACAGGTCAAGGAGAGAGCTGGGAAAGTGGCAGTCACTT 1965
Db 486 GluserIleuSerGlnuAsnArgSerArgIulGlnIleuGlyIyValGlySerGlnSer 505
QY 1966 AGCTTTTCGGGAGCATGAAATCATTTGAGGTCTCC 2001
Db 506 SerPheSerGlySerMetGlnIleIleGlnValSer 517

RESULT 2
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0


```

; SEQ ID NO 2
; LENGTH: 665
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-816-494-2

Alignment Scores:
Pred. No.: 0          Length: 665
Score: 472.00        Matches: 472
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 70.76%      Indels: 0
DB: 9                  Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-09-816-494-2 (1-665)

QY 586 ACCGTGCAAGAGCTGACTTATCCCGAGTCTCATTTCTGCGTGTGCTGAATGAC 645
   |||
Db 194 ThrCysProLysProAspPheIleProGluSerHisPheLeuAlaGValProValAsnAsp 213

QY 646 AGCTTTTGTGAGAAAATTTTCCGCGTGTGAGACAATCGATGATTTCATTGAGAAACA 705
   |||
Db 214 SerPheCysGluLysIleLeuProTyrLeuAspLysSerValAspPheIleGluLysAla 233

QY 706 AAGGCTCAATGATGTTGTCTAGTGCATGTTAGTGGGATCTCCGCGTCCGCGAC 765
   |||
Db 234 LysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThr 253

QY 766 ATCGTATCGCTACATCATGAAGAGATGACATGCTTTTAGATGAGCTTACAGATT 825
   |||
Db 254 IleAlaIleAlaIleTyrIleMetCysArgMetAspMetSerLeuAspGluAlaTyrArgPhe 273

QY 826 GTGAAAGAAAAGACCTTACTATATCTCAAACTTAATTTTGTGGCCAACTCTGAGAC 885
   |||
Db 274 ValLysGluLysArgProThrIleSerProAsnPheAsnPheLeuGlyGluLeuAsp 293

QY 886 TTGTGAGAGAGATTAGAACCAGATGAGATCAAGGCGCAAGAGCAAACTCAAGCTG 945
   |||
Db 294 TyrGluLysLysIleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeu 313

QY 946 CTGCACTTGAGAGAGCAATGAACTGTCCCTGCTGTCTAGAGGGTGAAGAAAGC 1005
   |||
Db 314 LeuHisLeuGluLysProAsnGluProValProAlaValSerGluGlyGluLysSer 333

QY 1006 GAGAGCGCCCTCAGTCCACCTCTGTGCGCACTGTCTAAGAGGCGAGGCAAGG 1065
   |||
Db 334 GluThrProLysSerProProCysAlaAspSerAlaThrSerGluAlaGlyGlnArg 353

QY 1066 CCGCGCAATCCGCGCAGCGTCCCGACCGTCCAGGCTCAGCCGCTGTTAAGAGAC 1125
   |||
Db 354 ProValHisProAlaSerValProSerValProSerValGlnProSerLeuGluAsp 373

QY 1126 ACCCGCGTGTGACAGGCGCTCAGTGGGCTGACCTTCCGCGACAGCGTGAAGACAGC 1185
   |||
Db 374 SerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSer 393

QY 1186 AATTAAGCTCAAGGTTCTCTCTCTGTGAATCAAACTGATTATTCAGCCAGCATG 1245
   |||
Db 394 AsnLysLeuLysArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMet 413

QY 1246 GCAGCATCTTTCATGGCTTCTCCCTCATGAGAAAGCTTGGAAATCAAAACCTTCC 1305
   |||
Db 414 AlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGluLysTyrLysProSer 433

QY 1306 ACTACTCTGGATGGAGCAACAAGCTATGCAAGTTCTCCCTGTTGAGAACTATCGAG 1365
   |||
Db 434 ThrThrLeuAspGlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGlu 453

QY 1366 CAGACTCTCCGAAAACAAGTCTGTATTAAGAGAAAGCAAGATCCCAAGAACTGCAGACC 1425
   |||
Db 454 GlnThrProGluThrSerProAspPheIleProGluSerHisPheLeuAlaGValProValAsnAsp 473

QY 1426 GCAGGCGCTTCAGACAGCAAGCAAGCAATTCGATTGCGTCAAGAACAGCAAGCTGCGC 1485
```

```

Db 474 AlaArgProSerAspSerGlnSerLysArgLeuHisSerValArgThrSerSerGly 493

QY 1486 ACCGCCAGAGAGTCCCTTTTATCTCCACTGCATCGAAGTGGAGCGGTGAGACAATTAC 1545
   |||
Db 494 ThrAlaGlnArgSerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyr 513

QY 1546 CACACAGCTTCCTTTTGGCGCTTCCACAGCCAGCAGCAGCTCAACGAAGCTGCTGCGC 1605
   |||
Db 514 HisThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGly 533

QY 1606 CTGGGCTTTAAGGGGTGGCACTCGGATATCTTGAGCCCCCAGACCTTAACCCCTTCCCTG 1665
   |||
Db 534 LeuGlyLeuLysGlyTyrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeu 553

QY 1666 ACCAGCAGCTGATTTTTCACAGAGTCTTCAACTTACTGCTGCTCAGCCATCTAC 1725
   |||
Db 554 ThrSerSerTyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyr 573

QY 1726 GGAGCGCAGTCCAGTACTCTGCGCTTACAGCTCCAGCCAGCTGCCACTTGCGGAGCA 1785
   |||
Db 574 GlyLysSerAlaSerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGln 593

QY 1786 GCTATTCGTGTCGAGCGCGAGAGCAAGCAAGTGAACAGCTGACTCGCGCGAGCTGG 1845
   |||
Db 594 ValTyrSerValArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTyr 613

QY 1846 CATGAAGAGAGCCCTTTGAAAAGCAATTAAACGACAGAGTGCCTCAATGAATTGGA 1905
   |||
Db 614 HisGluLysSerProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGly 633

QY 1906 GAGACATCATATTCAGAAACAGGTCACGGGAGAGCTGGGGAAGTGGCACTGACT 1965
   |||
Db 634 GluSerIleMetSerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSer 653

QY 1966 AGCTTTTGGCGCAGCATGGAATCATTTGAGGTCTCC 2001
   |||
Db 654 SerPheSerGlySerMetGluIleIleGluValSer 665

RESULT 3
US-09-964-277-2
; Sequence 2, Application US//09964277
; Patent No. US2002037170A1
; GENERAL INFORMATION:
; APPLICANT: Lucne, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125_434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-964-277-2

Alignment Scores:
Pred. No.: 0          Length: 665
Score: 472.00        Matches: 472
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 70.76%      Indels: 0
DB: 9                  Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-09-964-277-2 (1-665)

QY 586 ACCGTGCAAGAGCTGACTTATCCCGAGTCTCATTTCTGCGTGTGCTGAATGAC 645
   |||
Db 194 ThrCysProLysProAspPheIleProGluSerHisPheLeuAlaGValProValAsnAsp 213

QY 646 AGCTTTTGTGAGAAAATTTTCCGCGTGTGAGACAATCGATGATTTCATTGAGAAACA 705
```

Db 214 SerPheCysGluIuylIleuProTrpIeuAspIysSerValAspPheIleGluIuYsAla 233
QY 706 AAAGCCTCCAAATGATNGTCTTACTGACCTTTAGCTGGGATCTCCGCTCCGCAAC 765
Db 234 IysAlaSerAsnGlyCysValIleuValHisCysLeuAlaGlyIleSerArgSerAlaThr 253
QY 766 ATGCTATCGCCTACATCATGAAAGATGACATGCTCTTATAGATGAAGCTTACGATTT 825
Db 254 IleAlaIleAlaIyIleMerIysArgMetAspSerIleuAspGluAlaIyArgPhe 273
QY 826 GTGAAGAAAAAAGCCTTACTATATCTCAAACTTCAATTTTGGGGCAACTTCCTGAC 885
Db 274 ValIysGluIuYsArgProThrIleSerProAsnPheAsnPheLeuGlyGluIleuLeuAsp 293
QY 886 TATGAGAAGAATTAAGAACACAGCTGAGCATCAGGGCCCAAGACCAAACTTAAGCTG 945
Db 294 TyrGluIuYsIleIuYsAsnGlnThrGlyAlaSerGlyProIuYsSerIuYsLeu 313
QY 946 CTGCACCTGGAAGAACCAATGAACCTGTCCTGCTGCTCAGAGGGTGGACAGAAAG 1005
Db 314 LeuHisIleuGluIuYsProAsnGluProValProAlaValSerGluGlyGluIuYsSer 333
QY 1006 GAGAGCCCCCTCAGTCCACCTGTCGGACTCTGCTTCTCAGAGGACGACGACCAAGG 1065
Db 334 GluThrProIeuSerProProCysAlaAspSerAlaThrSerGluAlaIaGlyGlnArg 353
QY 1066 CCGGTGACCCCGCAGCGTGCAGCGGTGCCAGCGTGCAGCGTGCAGTGAAGAGAC 1125
Db 354 ProValHisProAlaSerValProSerValProSerValGlnProSerIleuLeuGluAsp 373
QY 1126 AGCCCGGTGTATACAGGCGCTCAGTGGGCTCAGCCTGTCGACAGACAGCTGAAGACAC 1185
Db 374 SerProIeuValGlnAlaIeuSerGlyLeuHisIleuSerAlaAspArgIeuGluAspSer 393
QY 1186 AATAAGCTCAAGCGTCTCTCTCTCTGATATCAAAATCAATTTCATTCAGCCAGCAG 1245
Db 394 AsnIuYsIeuYsArgSerPheSerIleuAspIleuYsSerValSerIuYsSerAlaSerMet 413
QY 1246 GCAGCATCTTACATGGCTTCTCTCATCAGAAAGATGCTTTGAAATATACAAACCTTCC 1305
Db 414 AlaAlaSerIeuHisGlyPheSerSerSerGluAspAlaIeuGluIuYrIyTrpProSer 433
QY 1306 ACTACTGTGATGGAGCAACAAGACATAGCCAGTCTCCCGCTTCAAGAACTATGAGAG 1365
Db 434 ThrThrIeuAspIeYsIuYsAsnIuYsCysGlnPheSerProValGlnIuYsIeuSerGlu 453
QY 1366 CAGACTCCCGAAACAGTCCGTATAGAGAGAACAGCATCCCAAGAAAGCTTGAGAGAC 1425
Db 454 GlnThrProGluThrSerProAspIuYsGluGluAlaSerIleProIuYsIeuGlnThr 473
QY 1426 GCCAGGCTTACAGACGACCAAGCAAGCATTCGATTCGATCAGAACACAGCAGCAGTGGC 1485
Db 474 AlaArgProSerAspSerGlnSerIuYsArgIeuHisSerValArgThrSerSerIeYs 493
QY 1486 ACCGCCAAGAGGCTCTTTATCTCCACATCCAGAGTGGAGAGCTGAGAGCAATTAC 1545
Db 494 ThrAlaGlnArgSerIeuIeuSerProIeuHisArgSerGlySerValGluAspAsnIuYr 513
QY 1546 CACACAGCTTCCTTTTCGCTTTCACACAGCAGCAGCACTTCAAGCAAGCTGCTGGC 1605
Db 514 HisThrSerPheIeuPheGlyIeuSerThrSerGlnGlnHisIleuThrIuYsSerAlaGly 533
QY 1606 CTGGGCTTAAAGGCTGGACATCGCATATCTTGGCCCCCAGACCTTACCTTCCCTG 1665
Db 534 IeuGlyIeuYsGlyTrpHisSerAspIleuAlaProGlnThrSerThrProSerIeu 553
QY 1666 ACCACAGCTGTATTTTGGCACAGAGTCCACACTTCTACTGCTCGCTCAGCATATAC 1725
Db 554 ThrSerSerTrpIyPheAlaThrGluSerSerHisPheIuYsSerAlaIleAlaIuYr 573
QY 1726 GAGGACAGTGCAGTTACTTGTGCTTACAGCTGACGCCAGCTGCCACTTGGAGAGACCA 1785

Db 574 GlyGlySerAlaSerTySerAlaTySerCysSerGlnIeuProThrCysGlyAspGln 593
QY 1786 GTCTATTCTGTGTGGCAGCGGACGAGCAAGCAAGTGAACAGAGCTGACCTGGGAGGCTGG 1845
Db 594 ValTySerValArgArgArgGlnIuYsProSerAspAlaAspSerArgArgSerTrp 613
QY 1846 CATGAAGAGACCCCTTTGAAAAAGCATTTAAAGCCAGAAAGCTGCCAAATGGAATTTGA 1905
Db 614 HisGluIuYsSerProPheGluIuYsGlnPheIuYsArgArgSerCysGlnMetGluPheGly 633
QY 1906 GAGAGCATCATTCAGAGAACAGTCAAGGGAAGCTGGGGAAGTGGAGTGCAGTCACT 1965
Db 634 GluSerIleMetSerGluAsnArgSerArgGluIuYsValGlySerGlnSer 653
QY 1966 AGCTTTGGGCGAGCATGGAATCATTTGAGGCTTCC 2001
Db 654 SerPheSerGlySerMetGluIleIleGluValSer 665
RESULT 4
US-10-072-012-680
Sequence 680, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zethusen, Bryan
APPLICANT: Patnrajan, Meera
APPLICANT: Shmukets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esna
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OR INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072, 012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265, 102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265, 514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265, 517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265, 412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265, 395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266, 406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266, 767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267, 057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266, 975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267, 459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 680
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-680

Alignment Scores:
Pred. No.: 0          Length: 665
Score: 472.00        Matches: 472
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 70.76%          Indels: 0
DB: 12                  Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-10-072-012-680 (1-665)

QY 586 ACCGTGCAAGAGCTGATCTTATCCCGAGTCTCATTTCTCGGTGCTGGAATGAC 645
    |||
DB 194 ThrCybProluSerProAmpRheileProGluSerHisPheLeuAlaGValProValaAenAmp 213

QY 646 AGCTTTTGAGAAATTTTCCCGTGTGAGCAAAATCGATGATTCTTGAAGAAACA 705
    |||
DB 214 SerPheCybGluLeuLeuProTyrLeuAmpLysSerValAmpRheileGluLysAla 233

QY 706 AAGGCTCAATGATGTGTTCTAGTGCATGTTTGTAGCTGGGATCTCCGCTCCGAC 765
    |||
DB 234 LysAlaSerAenGlyCybValLeuValHisCybLeuAlaGlyLysSerAenAlaThr 253

QY 766 ATCGCTATCGCTACATCATGAAAGAGATGACATGCTTTAGATGAGCTTACAGATT 825
    |||
DB 254 IleAlaIleAlaIleTyrIleMetLysArgMetAmpMetSerLeuAmpGluAlaTyrArgPhe 273

QY 826 GTGAAGAAAGAAAGCTTATATCTCCAACTTCAATTTTGTGGCCAACTCTCGAAC 885
    |||
DB 274 ValLysGluLysArgProThrIleSerProAenPheAenPheLeuLysGluLeuAenAmp 293

QY 886 TATGAGAGAGAGTTTAAAGACGAGTGCAGGAGTCAAGGGCCAAAGCAACTCAGCTG 945
    |||
DB 294 TyrGluLysLysLysLeuAenGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeu 313

QY 946 CTGCACTGAGAGAGCAATGAACTGTCCTCTGCTGTCTCAGAGGGTGAGCAAGAAAGC 1005
    |||
DB 314 LeuHisLeuGluLysProAenGluProValProAlaValSerGluGlyGlnLysSer 333

QY 1006 GAGAGCCCTCATGTCACCTGCTGCGACTGTGCTACTCAGAGGCAAGCAAGCAAG 1065
    |||
DB 334 GluThrProLysSerProProCybAlaAmpSerAlaThrSerGluAlaGlyGlnArg 353

QY 1066 CCGGTGCATCCGCGCAGCGGCCAGCGGCCAGGGGTGAGCGCGCTTGAAGAGAC 1125
    |||
DB 354 ProValHisProAlaSerAlaProSerAlaProSerAlaGlnProSerLeuGluAenAmp 373

QY 1126 AGCCGCTGATCAGAGCGCTCAGTGGGCTGACCTTCCGCAAGACAGGCTGAAAGACAC 1185
    |||
DB 374 SerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAmpArgLeuGluAenAmpSer 393

QY 1186 AATAAGCTCAAGCGTTCTTCTCTGTGATATCAAAATCAATTTCAATATTCAGCCACATG 1245
    |||
DB 394 AenLysLeuLysArgSerPheSerLeuAmpLysLysSerValSerLysSerAlaSerMet 413

QY 1246 GCGAGCATCTTACATGGCTTCTCTCATGAGAGAGCTTGAAGAACTTGAAGAACTTCC 1305
    |||
DB 414 AlaAlaSerLeuHisGlyPheSerSerSerGlnAmpAlaLeuLysTyrLysProSer 433

QY 1306 ACTACTCTGGATGGAGCAACAAGATATGCGCACTTCCCTGTTGAGAACTATCGAG 1365
    |||
DB 434 ThrThrLeuAmpGlyLysThrAenLysLeuCybGlnPheSerProValGlnGluLeuSerGlu 453

QY 1366 CAGACTTCGGAAGCAAGCTCTGTATTAAGAGAGAAAGCAAGCATCCCAAGAAAGTGCAGACC 1425
    |||
DB 454 GlnThrProGluLysThrSerProAmpLysGluGluAlaSerLysProLysLysLeuGlnThr 473

QY 1426 GCGAGGCTTTCAGACAGCCAGCAAGCATTTGCAATTCGCTCAGAACCCAGCAAGCATGGCC 1485

```

```

DB 474 AlaArgProSerAmpSerGlnSerLysArgLeuHisSerValArgThrSerSerGly 493

QY 1486 ACCGCCCAAGAGTCCCTTTTATCTCCACTGCATCCGAGTGGAGCGCTGAGCAATTCAC 1545
    |||
DB 494 ThrAlaGlnArgSerLeuSerProLeuHisArgSerGlySerValGluAenAenTyr 513

QY 1546 CACACAGCTTCTTTTGGGCTTTCCACAGCCAGGACGACCTCAAGCAAGCTGCTGGCC 1605
    |||
DB 514 HisThrSerPheLeuPheGlyLeuSerThrSerGlnHisLeuThrLysSerAlaGly 533

QY 1606 CTGGGCTTAAAGGCTGCGACCTCGGATATCTTGAGCCCGCCAGACCTTACCCCTTCCCTG 1665
    |||
DB 534 LeuGlyLeuLysGlyTyrHisSerAmpLysAlaProGlnThrSerThrProSerLeu 553

QY 1666 ACCAGAGCTGATTTTCCACAGAGTCTTCAACTTACTGCTTCCAGCATTCAC 1725
    |||
DB 554 ThrSerSerTyrTyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyr 573

QY 1726 GAGAGCAGTCCAGTACTCTGCTTACAGCTGACAGCTGCGACCTGCGGAGACCA 1785
    |||
DB 574 GlyLysSerAlaSerLysSerAlaTyrSerCybSerGlnLeuProThrCybGlyAmpGln 593

QY 1786 GTCTATTCTGTCGAGCGGCGAGCAAGCCAACTGACAGACTGACTCGCGCGGAGCTG 1845
    |||
DB 594 ValTyrSerValArgArgArgGlnLysProSerAmpArgAlaAmpSerArgArgSerTyrP 613

QY 1846 CATGAGAGAGCCCTTTGAAAGAGATTTAAACGAGAGCTGCCAAATGGAATTTGCA 1905
    |||
DB 614 HisGluGluSerProPheGluLysGlnPheLysArgArgSerCybGlnMetGluPheGly 633

QY 1906 GAGAGCATATATTCAGAGAAAGGTCACGGGAAAGCTGGGAAAGTGGGCACTGACT 1965
    |||
DB 634 GluSerLysMetSerGluAenArgSerArgGluGluLeuGlyLysValGlySerGlnSer 653

QY 1966 ACCTTTTCCGCGAGCATGGAATTCATTGAGTCTCC 2001
    |||
DB 654 SerPheSerGlySerMetGluIleIleGluValSer 665

RESULT 5
US-10-072-012-681
; Sequence 681, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patcurajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Bha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Raselli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30

```

```
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 60/266,767
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: 60/267,057
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/266,975
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/267,459
/ PRIOR FILING DATE: 2001-02-08
/ Remaining Prior Application data removed - See file wrapper or PAM.
/ NUMBER OF SEQ ID NOS: 1391
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 681
/ LENGTH: 665
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-072-012-681
```

Alignment Scores:

```
Pred. No.: 0 Length: 665
Score: 472.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.76% Indels: 0
DB: 12 Gaps: 0
```

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-072-012-681 (1-665)

```
QY 586 ACCCTGCAAAAGCTGACTTTATCCCGAGTCTCATTTCTCGTGCTGCAATGAC 645
DB 194 ThrCysProIySerProAspHeIleProGluSerHisPheLeuAlaGlyValProAlaIasnAsp 213
QY 646 AGCTTTTGAGAAAATTTTGGCGGTGGTGAACAATCGATGATTCTTGAAGAAAGA 705
DB 214 SerPheCysGluIlySerIleLeuProIleuAspIlySerValAspPheIleGluIlySala 233
QY 706 AAAGCTCAATGATGATGTTCTAGTCACTGTTTAGCTGGATCTCCCGCTCCGCCACC 765
DB 234 LysAlaSerAsnGlyCysValIleuValHisCysLeuAlaGlyIleSerArgSerAlaThr 253
QY 766 ATCGCTATCGCTACATCATGAAGAGATGACATGCTTTAGATGAAGCTTACAGATT 825
DB 254 IleAlaIleAlaIlyIleMetIlySerArgMetAspMetSerLeuAspGluAlaIlyArgPhe 273
QY 826 GTGAAAAGAAAAAGACCTNACTATATCTCAAACTTCAATTTTCTGCGCCAACTCTGAGC 885
DB 274 ValIlySerGluIlyArgProThrIleSerProAsnPheAsnPheLeuGlyGluIleuIasnAsp 293
QY 886 TATGAGAGAAATTAAGAACAGACCTGAGATGACAGCCCAAAAGCAAAATCTCAAGCTG 945
DB 294 TyrGluIlySlyIleIlyAsnGlnIthrGlyAlaSerGlyProIlySerIlySleuIlySleu 313
QY 946 CTGCACTGAGAGAACCAATGAACTGCTGCTGCTGCTGCTGAGAGGAGTGCAGAAAAAGC 1005
DB 314 LeuHisIleuGluIlyPheAsnGluProAlaProAlaValSerGluGlyGlyIlyIlySer 333
QY 1006 GAGAGCGCCCTCAGTCCACCTCTGCGCCCACTCTGCTTACCTCAGAGCAGCAAGCAAGG 1065
DB 334 GluThrProIySerProProCysAlaIasnSerAlaIthrSerGluAlaIalGlyIlyIlyArg 353
QY 1066 CCGCGCATCCCGCCAGAGGTCGCCAGCGTGGCCAGCGTGCAGCGCTCGCTGTTAAGAGAC 1125
DB 354 ProAlaHisProAlaSerValProSerValProSerValGlnProSerIleuIleuIlySleu 373
```

```
QY 1126 AGCCCGCTGTACAGAGCGCTCAGTGGCTGACCTGTCCGAGACAGGCTGAGAACAGC 1185
DB 374 SerProIleuValGlnAlaIleuSerIlyLeuHisIleuSerAlaAspArgIleuIlySleu 393
QY 1186 AATTAAGCTCAAGCGTTCTCTCTGTGATATCAAAATCAGTTTCATTTATTCAGCCAGCATG 1245
DB 394 AsnIlySleuIlyArgSerPheSerIleuAspIleIlySerValSerIlySerAlaSerMet 413
QY 1246 GGAGCATCTTAATGAGGCTTCTCCCTCATCAGAAAGTGTCTTTGGAATCTATCAAACTTCC 1305
DB 414 AlaAlaSerLeuHisGlyIlyPheSerSerSerGluAspAlaLeuGluIlyIlyIlyProSer 433
QY 1306 ACTACTGTGATGGAGCAACAAGATATGCAAGTCTCCCTGTTCAAGAACTATACGAGAG 1365
DB 434 ThrThrIleuAspGlyIlyThrAsnIlySleuIlySerGlnPheSerProValGlnGluIleuSerGlu 453
QY 1366 CAGACTCCGAAACCAAGTCTGTATAGAGAGAGAGCCAGCATCTCCCAAGACGTGCAGAC 1425
DB 454 GlnThrProGluIlyThrSerProAspIlySerGluIlyAlaSerIleProIlyIlySleuGlnThr 473
QY 1426 GCCAGCGCTTCAGACAGCAGCAGAGGATTTGCTTGGTTCAGAAACCAAGCAGCAGTGGC 1485
DB 474 AlaArgProSerAspSerGlnSerIlySerIlyAsnIlySerValAlaGlyThrSerSerIly 493
QY 1486 ACCGCCAGAGGTCCTTTTATCTTCACTGATCGATCGAAGTGGAGGCGTGAAGCAATTAC 1545
DB 494 ThrAlaGlnArgSerIleuSerProIleuHisArgSerGlySerValGluAspAsnIly 513
QY 1546 CACACAGCTTCTCTTTTGGCGCTTTCCACCAAGCAGCAGCAGCAGCAGTCTGCTGGC 1605
DB 514 HisThrSerPheLeuPheGlyIlyLeuSerThrSerGlnGlnHisIleuThrIlySerAlaIly 533
QY 1606 CTGGCGCTTAAGGGGTGCGACATCGGATATCTTGGCGCCCGCCAGACTTCACCTTCCCTG 1665
DB 534 LeuGlyLeuIlySlyIlyThrPheIlySerAspIleuAlaProGlnIlyThrProSerIleu 553
QY 1666 ACCAGCAGCTGATTTTGGCAGCAGAGTCTCACACTTCTCTGCTGCTCAGCCATCTAC 1725
DB 554 ThrSerSerIlyPheAlaIlyThrGluSerSerHisPheIlySerAlaIleIly 573
QY 1726 GAGGCGAGTGCAGTACTCTTGTCTTACAGCTGCAGCAGCAGCTGCCACTTGGGAGACCA 1785
DB 574 GlyIlySerAlaSerIlySerAlaIlySerCysSerGlnIleuProIlyCysGlyAspGln 593
QY 1786 GTCTATCTGTGCCAGCGCGCAGAAAGCAAGTCAAGAGTGAAGTCTGCGCGGAGACTGG 1845
DB 594 ValIlySerValAlaArgAlaArgGlnIlyProSerAspArgAlaAspSerAlaArgSerIly 613
QY 1846 CATGAAGAGAGCCCTTTGAAGAAAGCAATTAAAGCAGAGCTGCCAAATGGAATTTGGA 1905
DB 614 HisGluIlySerProPheGluIlyGlnPheIlyArgArgSerCysGlnMetGluPheGly 633
QY 1906 GAGAGCATCATGTGAGAGAACAGGTCAAGGAGAGAGCTGGGAGAGTGGGAGTCAAGTCT 1965
DB 634 GluSerIleMetSerGluAsnArgSerArgGluIlyIleuGlyIlyValGlySerGlnSer 653
QY 1966 AGCTTTCCGGCAGACATGAAATCATTTGAGAGTCTCC 2001
DB 654 SerPheSerGlySerMetGluIleIleGluValSer 665
```

RESULT 6
US-10-168-506-14
Sequence 14, Application US/10168506
Publication No. US20040053229A1
GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY D.
APPLICANT: MARTINEZ, RICARDO
APPLICANT: WHYTE, DAVID
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
APPLICANT: HILL, RON
APPLICANT: FLANAGAN, PETER
TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES

```

FILE REFERENCE: 038602/1351
CURRENT APPLICATION NUMBER: US/10/168,506
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: PCT/US00/34736
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-10-168-506-14

Alignment Scores:
Pred. No.: 0 Length: 665
Score: 472.00 Matches: 665
Percent Similarity: 100.00% Conservative: 472
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.76% Indels: 0
DB: 12 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-168-506-14 (1-665)

QY 586 ACCTGTCGCAAGCCGACTTATCCCGAGTCTCATTTCCGCGGTGCGCTGTAATGAC 645
DB 194 ThrcyspPolyspProaspPheIIleProGlnSerHisPheLeuArgValProValAsnsp 213
QY 646 AGCTTTGTGAGAAATTTTCCGCTGTTGGACAATGACATGATGATTGATGAGAAACA 705
DB 214 SerPheCysGlnIuylsIleuProItrPheuaspIysSerValAspPheIIleGlnIuylsAla 233
QY 706 AAAGCTTCATGATGATGTGTTCTAGTGACATGTTTACTGGGATCTCCCGCTCCGACCC 765
DB 234 LysAlaSerAsnGlnCysValLeuValHisCysLeuAlaGlnIylIleSerArgSerAlaThr 253
QY 766 ATCGCTATCGCTCATGATGATGAAGAGATGAGACATGTTTATGATGATGATGATGATTT 825
DB 254 ILeaIleAlaIleAlaIylIleMeclLysArgMetAspSerLeuAspGlnAlaIleArgPhe 273
QY 826 GTGAAGAAAAAAGACCTTACTATATCTCCAAACTTCAATTTTCTGGCCAACTCCTGGAC 885
DB 274 ValIysGlnIuylsArgProThrIIleSerProAsnPheAsnPheLeuGlnIleuLeuAsp 293
QY 886 TATGGAAGAAGATTAGAACCCAGCTGGAGCATGAGGCGCAAGACCAACTCAAGCTG 945
DB 294 TyrGlnIuylsIleIysAsnGlnThrGlnIylAsSerGlnIylPolySerIysLeuIysIleu 313
QY 946 CTGACACCTGGAGAAAGCCAAATGAACCTTCCCTGCTGCTCTAGAGGGTGGACAGAAAAGC 1005
DB 314 LeuHisIleuGlnIuylsProAsnGlnProValProAlaValSerGlnIuylGlnIysSer 333
QY 1006 GAGACGCCCTCAGTCCAGCCACCTGTGTCGACCTCTGTACTCTGAGAGGCGAGGACAAAGG 1065
DB 334 GlnItrProIleuSerProIleuProCysAlaAspSerAlaItrSerGlnAlaIleGlnItrG 353
QY 1066 CCCGTGATCCCGCCAGCGTGCACAGCTGCCAGCTGCAGCGCTGCTGTAGAGAC 1125
DB 354 ProValHisProAlaSerValProSerValProSerValGlnProSerLeuLeuGlnIuAsp 373
QY 1126 AGCCCGCTGTGATCAGGCCCTCAGTGGGCTGCACCTGTCCGACAGAGGCTGGAAGACAGC 1185
DB 374 SerProIleuValGlnAlaLeuSerIylLeuHisIleuSerAlaAspArgLeuGlnIuAspSer 393
QY 1186 AATAAGCTCAGACCGTCTCTCTCTGTGATATCAATCAGTTTCAATTGACGAGCATG 1245
DB 394 AsnIuylsIleuIysArgSerPheSerIleuAspIleIylsSerValSerTyrSerAlaSerMet 413
QY 1246 GCAGCATCTTACATGAGCTTCTCTCATCAGAAAGATGCTTTGATATACAAACTTCC 1305
DB 414 AlaIleSerIleuHisGlnIylPheSerSerSerGlnAlaIleuGlnIuIylTyrIylProSer 433
QY 1306 ACTACTCTGGATGGGACCAACAGTATGCCAGTTTCCCTGTTCAGAAACTATCGAG 1365

```

Dd	434	ThrThrIleuAspGlyThrThrAsnLysLLeuCYGlnPheSerProValGlnGluLeuSerGlu	455
Oy	1366	CAGACTCCCGAAGACAGTCTGATAGAGAGAGAACCGACATCCCGAAGAGAGCTGCAGACC	1425
Dd	454	GlnThrProGluThrSerProAspLysGlnGluValaSerIleProLysLysLLeuGlnThr	473
Oy	1426	GCCAGGCCCTTCAGACAGCCAGAGACCGATTCGCGTTCAGAACACACAGAGAGGC	1485
Dd	474	AlaArgProSerAspSerGlnSerLysArgLysHisSerValArgThrSerSerSerGly	493
Oy	1486	ACCGCCAGAGAGTCCCTTTTATCTCCCACTCATCGAAGTGGAGCGCGAGAGCAATTAC	1545
Dd	494	ThrIaGlnArgSerLysLeuLeuSerProLysHisArgSerGlySerValGluAspAsnThr	513
Oy	1546	CACACAGACTTCCTTTTCGCGCTTTCACAGCCAGAGACCTTCAGCAAGTGTGCTGC	1605
Dd	514	HisThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGly	533
Oy	1606	CTGGGCGCTTAAGGCGCTGGCACTGGGAATCTTGAGCCCCCAAGACTTACCCTTCCTCG	1665
Dd	534	LeuLysLeuLysGlyLysPheLysSerAspLysLeuValProGlnThrSerThrProSerLeu	553
Oy	1666	ACCAAGAGCTGTGTATTTTGGCAACAGAGTCCCACTTCACTCTGCTGCAGCATCTAC	1725
Dd	554	ThrSerSerThrPlyrPheAlaThrGlnLysSerSerHisPheThrSerAlaSerAlaIleThr	573
Oy	1726	GGAGGACAGTGCAGATTACTGTGCTTCAGCTGCAGCTGCAGCCAGCTGCCCACTTGCGAGCAAA	1785
Dd	574	GlyGlySerAsnLysSerThrLysSerCysSerGlnLeuProThrCysGlyAspGln	593
Oy	1786	GTCATATTCGTGCGCAGCGCGCAGAGCCAAAGTACAGAGACTGACTCGCGCGGAGCTGG	1845
Dd	594	ValLysSerValArgArgArgGlnLysProSerAspArgLysAspSerArgArgSerThr	613
Oy	1846	CATGAAGAGAGCCCTTTGAAAAAGCAATTAAACGAGAAAGCTGCCAAATGSAATTGGA	1905
Dd	614	HisGlnLysSerProPheGlnLysGlnPheLysArgArgSerCysGlnMetGluPheGly	633
Oy	1906	GAGAGCATCATGTTCAGAAACAGGTCACGGGAAAGCTGGGGAAGTGGCGAGTCACTCT	1965
Dd	634	GlnSerIleMetSerGlnAsnArgSerArgGlnGluLeuGlyValGlySerGlnSer	653
Oy	1966	AGCTTTTCGGGCGAGCATGGAATCATTTGAGGTCTCC	2001
Dd	654	SerPheSerGlySerMetGlnIleIleGluValSer	665
RESULT 7			
US-10-343-357-7			
; Sequence 7, Application US/10343357			
; Publication No. US20040058341A1			
; GENERAL INFORMATION:			
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y.Tom			
; APPLICANT: ELLIOTT, Vicki S.; RAMKIDAR, Jayalaxmi			
; APPLICANT: YAO, Monique G.; BURFORD, Neil			
; APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.			
; APPLICANT: GANDHI, Ameena R.; ARVIZO, Chandra S.			
; APPLICANT: LEE, Ernestine A.; HARALALA, April J.A.			
; APPLICANT: LU, Dzung Aina M.; TRIBOULEY, Catherine M.			
; APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.			
; APPLICANT: YOE, Henry; WARREN, Bridget A.			
; APPLICANT: NGUYEN, Daniel B.; CHAWLA, Narinder K.			
; APPLICANT: KEARNEY, Liam			
; TITLE OF INVENTION: PROTEIN PHOSPHATASES			
; FILE REFERENCE: PI-0173 PCT			
; CURRENT APPLICATION NUMBER: US/10/343.357			
; CURRENT FILING DATE: 2003-01-28			
; PRIOR APPLICATION NUMBER: PCT/US01/23716			
; PRIOR FILING DATE: 2001-07-26			
; PRIOR APPLICATION NUMBER: US 60/221,679			
; PRIOR FILING DATE: 2000-07-28			
; PRIOR APPLICATION NUMBER: US 60/223,272			
; PRIOR FILING DATE: 2000-08-03			
; PRIOR APPLICATION NUMBER: US 60/224,309			

; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 60/226,728
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/229,254
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: US 60/231,366
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 665
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CD1
US-10-343-357-7

Alignment Scores:

Pred. No.: 0 Length: 665
Score: 472.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.76% Indels: 0
DB: 12 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-10-343-357-7 (1-665)

QY 586 ACCGTGTCAGAGCTGATTTATCCCGAGTCTCATTTCCGCGTGTGCTGTAATGAC 645
DB 194 TnCyProLyserProAspRheIleProGluSerHisPheLeuValProValaMet 213
QY 646 AGCTTTGAGAAATTTTGGCGGTGGAGCAAACTGATGATTTCTTGAAGAGA 705
DB 214 SerPheCysGluLysIleLeuProIleuAspLysSerValAspPheIleGluLysAla 233
QY 706 AAAGCTCAATGATGTGTTCTAGTGCATGTTTAGCTGGATCTCCGCGTCCGAC 765
DB 234 LysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThr 253
QY 766 ATGCGTATGCGCTACATCATGAAAGAGATGACATGTCTTTAGATGAAGCTTACAGATT 825
DB 254 IleAlaIleAlaIleArgIleMetLysArgMetAspMetSerLeuAspIleAlaIleArgPhe 273
QY 826 GTGAAGAAAGAAAGACCTATATCTCCAACTTCAATTTCTGGGCCAACTCCGAGAC 885
DB 274 ValLysGluLysArgProThrIleSerProAsnPheAsnPheLeuGlyGluLeuLeuAsp 293
QY 886 TATGAGAGAAGATTTAAGAACAGACTGAGCATGAGGCGCAAGAGCAAACTCAAGCTG 945
DB 294 TyrGluLysLysIleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeu 313
QY 946 CTGACCTGAGAGAGCCAAATGAACTGTGCTGCTGTCTCAGAGGGGTGAGCAAGAAAGC 1005
DB 314 LeuHisIleGluLysProAsnGluProValProAlaValSerIleGlyGlnLysSer 333
QY 1006 GAGAGCCCTCGATCCACCTGTCGCGACTGCTGACCTGAGGACGACAGGCAAGAAAG 1065
DB 334 GluThrProLeuSerProProCysAlaAspSerAlaThrSerIleAlaIleGlyGlnArg 353
QY 1066 CCGGTGATCCCGCAGCGTGCAGCGTGCAGCGGCGAGCGTGCAGCGCTGTGTAAGAGAC 1125
DB 354 ProValHisProAlaSerValProSerValProSerValGlnProSerIleuLeuGluAsp 373
QY 1126 AGCCGCTGTGATGAGCGCTCAGTGGGCTGACCTGTCCGAGACAGAGCTGGAAGACAGC 1185
DB 374 SerProLeuValGlnAlaLeuSerGlyLeuHisIleuSerAlaAspArgLeuGluAspSer 393
QY 1186 AATTAAGCTCAAGGCTCTTCTCTGAGATATCAATTCAGTTTCAATTCAGCAGACAG 1245
DB 394 AsnLysLeuLysArgSerPheSerIleuAspIleLysSerValSerTyrSerAlaSerMet 413
QY 1246 GCAGCATCTTACATGAGCTTCTCTCATGACAGAGATGCTTTGGAATATACAAACCTTCC 1305

DB 414 AlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSer 433
QY 1306 ACTACTCTGATGGAGCAACAGCTATGCCAGTTCTCCCTGTTCAAGAACTATCGAG 1365
DB 434 ThrThrLeuAspGlyThrAsnLysLeuGlyGlnPheSerProValGlnGluLeuSerGlu 453
QY 1366 CAGACTCCCGAAACAGTCTGTATGAGAGAGAGAGACATCCCAAGAGCTGAGAC 1425
DB 454 GlnThrProGluThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThr 473
QY 1426 GCCAGGCTTTCAGACAGCGCAAGCAAGCATTTGCAATTCGTCAGAACAGACAGCTGAC 1485
DB 474 AlaArgProSerAspSerGlnSerLysArgLeuHisSerValArgThrSerSerGly 493
QY 1486 ACCGCCCGAGAGTCCCTTTTATCTTCACATGCATGAAAGTGGAGCGGTGAGACAAATTAC 1545
DB 494 ThrAlaGlnArgSerLeuLeuSerProLeuHisAspSerGlySerValGluAspAsnTyr 513
QY 1546 CACACAGAGCTTCTTTCGACCTTTCACACAGCGCAGACACCTCAGCAAGCTCTGAC 1605
DB 514 HisThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGly 533
QY 1606 CTGGGCTTAAAGGCTGCGACCTCGGATATCTTGCGCCCGCCAGACCTTACCCCTTCCCTG 1665
DB 534 LeuGlyLeuLysGlyTyrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeu 553
QY 1666 ACCAGACAGCTGATATTTGCCACAGAGCTCCACACTTCTCTGCTGCGCTGACCTTAC 1725
DB 554 ThrSerSerTyrPyrPheAlaThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyr 573
QY 1726 GGAGCAGTGGCAGATTAATCTGCTTACAGCTGACAGCCAGCTGCGCCACTTGGCGAGACCA 1785
DB 574 GlyGlySerAlaSerTyrSerAlaIleTyrSerCysSerGlnLeuProThrCysGlyLysPheGln 593
QY 1786 GTTATTTGTGCGCAGCGCGAGAGCCAAAGTGAAGCAAGCTGACTCGCGCGAGCTGG 1845
DB 594 ValTyrSerValArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTyrP 613
QY 1846 CATGAAGAGAGCGCCCTTTGAAAGAGCAAGTAAAGCAAGCTGCAATGAAATTTGGA 1905
DB 614 HisGluGluSerProPheGluLysGlnPheLysArgLysSerCysGlnMetGluPheGly 633
QY 1906 GAGACATCATGTCTGAGAAACAGGTCAAGGAGAGAGCTGAGGAGAAAGTGGAGTCACT 1965
DB 634 GluSerIleMetSerGluAsnArgSerArgGluGluLeuGlyLysValGlySerIleSer 653
QY 1966 AGCTTTTGGGCGAGCATGAAATCATTAAGGTTCTCC 2001
DB 654 SerPheSerGlySerMetGluIleIleGluValSer 665

RESULT 8
US-10-094-749-2312
; Sequence 2312, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YORI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI

```

; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 08435/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 2312
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-094-749-2312

Alignment Scores:
Pred. No.: 0 Length: 665
Score: 472.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.76% Indels: 0
DB: 15 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-10-094-749-2312 (1-665)

QY 586 ACCTGTCCAAAGCCTGATCTTATCCCGAGCTCTCATTTCTGCGTGTGCTGTAATGAC 645
DB 194 ThTCyProlsProlsPheleProgluSerHisPheleuArgValProValaanaP 213
QY 646 AGCTTTTGTGAGAAATTTTGGCGTGTGGAACAATCAGTAATTTTCATGAGAAAGCA 705
DB 214 SerPheCySeGlulysIleuProTyrPleuAspLysSerValAspPheIleGlulysAla 233
QY 706 AAAGCCTCCCAATGATGTGTTTCTAGTCACTGTTTGTGAGTCCGCTCCGCGCAC 765
DB 234 LysAlaSerAsnGlyCyValIleuValHisCySeuAlaGlyIleSerAspSerAlaThr 253
QY 766 ATGCTATCGCTTACATCATGAAGAGATGACATGCTTTAGATGAAGCTTACAGATT 825
DB 254 IleAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPhe 273
QY 826 GTGAAGAAAGAAAGACTTACTATATCTCCAACTTCAATTTTCTGGCCCACTCTCGAC 885
DB 274 ValIleGlulysArgProThrIleSerProAspPheAsnPheLeuGlyIleuLeuAsp 293
QY 886 TATGAGAGAAGATTAAAGACAGACTGAGACATCAGGCGCAAGCAAACTCAAGCTG 945
DB 294 TyrGlulysIleuLysAsnGlnThrGlyAlaSerGlyProlysSerIleuLysLeu 313
QY 946 CTGCACCTGGAGAGCCAAATGAACCTGTCTCTCTCAGAGGGTGGACAGAAAGC 1005
DB 314 LeuHisIleuGlulysProlAsnGluProlValIleSerGluGlyIleuLysSer 333
QY 1006 GAGACGCCCTCAGTCCACCTGTGCCACTGTGCTACTCAGAGGACAGGACAAAG 1065
DB 334 GluThrProlSerProProCyAlaAspSerAlaThrSerGluAlaIleGlyIleuArg 353
QY 1066 CCGGTGATCCCGCGCGCGCCAGCGTCCAGCGTCCAGCGTCCGCTGTTAGAGGAC 1125
DB 354 ProValHisProlAspSerValProSerValProSerValGlnProSerLeuLeuGluAsp 373
QY 1126 AGCCCGCTGTACAGCGCTCAGTGGGCTGACCTGTCCGACAGAGAGCTGAGAGCAGC 1185
DB 374 SerProlSerValGlnAlaLeuSerGlyLeuHisIleuSerHisAspArgLeuGluAspSer 393
QY 1186 AATPAGCTCAAGCGTTCTTCTCTCTGAGATATCAATCAGTTTCATATTACGACGACG 1245
DB 394 AsnLysLeuLysArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMet 413
QY 1246 GCGACATCCTTACATGCTCTCTCTCTCATCAGAGAGCTTTGAGAAATCTCAACCTTCC 1305
```

```

DB 414 AlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSer 433
QY 1306 ACTACTCTGGATGGAGCAACAAAGCTATGCTCCCTGTCAGGAATCATGCGAG 1365
DB 434 ThrThrLeuAspGlyThrAsnLysLeuCyGlnPheSerProValGlnIleuSerGly 453
QY 1366 CAGACTCCCGAAACAGCTCTGTATAGAGAGAGAGCAGCATCCCAAGAGCTGACAGCC 1425
DB 454 GlnThrProlGluThrSerProlAspLysGluGluAlaSerIleProlLysIleuGlnThr 473
QY 1426 GCCAGGCTTTCAGACAGCCAGCAAGCAGATTCATTCGTCGTAACCAAGCAGCAGTGC 1485
DB 474 AlaArgProSerAspSerGlnSerIleLysArgLeuHisSerValArgThrSerSerGly 493
QY 1486 ACCGCCCAAGCTCCCTTTATCTCCACCTGCTCAAGAGTGGAGGAGCAATTC 1545
DB 494 ThrAlaGlnArgSerLeuSerProLeuHisArgSerGlySerValGluAspAsnTyr 513
QY 1546 CACACAGCTTCTTTTGGCTTTCACACAGCCAGCAGCAGCAGCAGTCTGCGGC 1605
DB 514 HisThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisIleuThrLysSerAlaGly 533
QY 1606 CTGGGCTTAAAGGCTGCGCAGCTCGATATCTTGCCCCCAGACCTTACCCCTTCCCTG 1665
DB 534 LeuGlyLeuLysGlyTyrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeu 553
QY 1666 ACCAGCAGCTGTATTTTCCACAGAGTCTTACACTTCTGCTGCTCAGCATCTAC 1725
DB 554 ThrSerSerTyrTyrPheAlaThrGluSerSerHisPheTyrSerLysSerAlaIleTyr 573
QY 1726 GGAGGAGAGCCAGTACTCTGCTTACAGCTCAGCAGCAGCTGCCACTTGCAGAGACAA 1785
DB 574 GlyGlySerAlaSerTyrSerAlaTyrSerCySeGlnLeuProThrCySeGlyAspGln 593
QY 1786 GTCTATTCTGTGCGAGCGCGGAGCAAGCCAGTGAACAGCTGACTCGCGGCGAGCTGG 1845
DB 594 ValTyrSerValArgArgArgGlnLysProlSerAspArgAlaAspSerArgArgSerTyr 613
QY 1846 CATGAAGAGAGCCCTTTGAAGAGAGCTTAAACGCAAGAGCTGCCAAATGGAATTGGA 1905
DB 614 HisGluGluSerProPheGluLysGlnPheLysArgArgSerCySeGlnMetGluPheGly 633
QY 1906 GAGAGCATCATGTCAGAGAACAGGTCAAGGAGAGCTGGGGAAGTGGGAGTCACTGCT 1965
DB 634 GluSerIleMetSerGluAsnArgSerArgGluGluLeuGlyLysValIleSerGlnSer 653
QY 1966 AGCTTTTCGGGAGCATGGAATCATTTGAGTCTCC 2001
DB 654 SerPheSerGlySerMetGluIleIleGluValSer 665

RESULT 9
US-10-377-072-26
; Sequence 26, Application US/10377072
; Publication No. US20040009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Myoung
; APPLICANT: Tsai, Feng-Ying
; TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULARS
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MP103-0180NMIM
; CURRENT APPLICATION NUMBER: US/10/377,072
; PRIOR FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
```



```
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; PRIOR FILING DATE: 2001-03-22
; Remaining prior application data removed - See file wrapper or PAM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 665
; TYPE: PKT
; ORGANISM: Homo Sapiens
; US-10-377-072-26

Alignment Scores:
pred. No.: 0          Length: 665
Score: 472.00        Matches: 472
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.76%          Indels: 0
DB: 15                  Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-10-377-072-26 (1-665)

QY 586 ACCTGTCAGAAAGCTGACTTATCCCGAGTCTCATTTCTCGTGGCTGTGAATGAC 645
   |||||
Db 194 ThrcyProlyseroAspheiIeProgluserHisPheleuAryGAlProValaanaP 213

QY 646 AGCTTTTGTGAGAAATTTTGGCCGTGGTGACAAATCAGTAGATTTCATTGAGAAAGA 705
   |||||
Db 214 SerPheCysgluYsIleleuProTlrleuAsplySerValAspheiIeGluYsAla 233

QY 706 AAAGCTCCAGTAGGTGTTCAGTGCACGTGTTGCTGGAGATCCCGCTCCGCGAC 765
   |||||
Db 234 LysAlaSerAsnGlySvalleuValHisCysleuAlaGlyIleSerIrgSerAlaInr 253

QY 766 ATGCTATGCGCTCATCATGAAGAGATGACATGCTCTTAGATGAACTTACAGATT 825
   |||||
Db 254 IleAlaIleAlaIleYrIleMetLysArgMetAspMetSerleuAspGluAlaIYrArgPhe 273

QY 826 GTGAAAGAAAAAGACTTATATCTCAAACTTAAATTTTGTGGCCAACTCTGAGAC 885
   |||||
Db 274 ValIySgluYsIrgProThrIleSerProAsnPhaAsnPhaIleuYlGluIleuAsp 293

QY 886 TATGAGAAAGATTAGAACAGACATGAGACATCGGGCCAAAGGCAAACTCAAGCTG 945
   |||||
Db 294 IYrGluYlYsIleYsAsnGlnInrGlyAlaSerGlyProIySerIySleuYsIleu 313

QY 946 CTGCACTGAGAGAGCAATGAACTGTCCCTGCTGTCTGACAGGGGTGAGAGAAAGC 1005
   |||||
Db 314 leuHisIleuGluYsIrgProAsnGlnIurProAlaValaSerGluYlGlnIySsIer 333

QY 1006 GAGAGCCCCCTGACCTGCTGTGGCGACTGTCTACTCTCAAGGACAGAGCAAGAG 1065
   |||||
Db 334 GluInrProIeuserProProCysAlaAspSerAlaThrSerGluAlaIaGlyInArg 353

QY 1066 CCGCGCATCCCGCGAGCGGTGCCAGCGTGGCCAGCGTGCAGCGCTGTGTAAGAGC 1125
   |||||
Db 354 ProValHisIProAlaSerValProSerValProSerValGlnProSerIleuGluuAsp 373

QY 1126 AGCCCGCTGTACAGGCGCTCAGTGGGCTGACCTGTCCGACAGACAGGCTGAGAACAGC 1185
```

```
|||||
Db 374 SerProleuValGlnAlaIeuserGlyleuHisIeuserAlaAspArgLeuGluuAspSer 393

QY 1186 AATAAGCTCAAGCGCTCTCTCTCTGATATCAAAATGATTTCAATATTCAGCCGACATG 1245
   |||||
Db 394 AsnIySleuYsIrgSerPheSerleuAspIleYsIerValSerIySerAlaSerMet 413

QY 1246 GCAGATCCTTACATAGGCTCTCCCTCATCAGAAAGCTTTGGAATATCTAACAACTTCC 1305
   |||||
Db 414 AlaAlaSerleuHisGlyPheSerSerSerGluuAspAlaIeugluYrIYrIySroSer 433

QY 1306 ACTACTGTGATGGAGCAACAAGCTATGCCAGTTCTCCCTGTTCAGAACTATGGAG 1365
   |||||
Db 434 ThrThrIleuAspGlyThrAsnIySleuYsGlnPheSerProValGlnIuIeuserGlu 453

QY 1366 CAGATCCCGAAACCAAGTCTGATTAAGAGAGAGACCAATCCCAAAAGCTGAGACC 1425
   |||||
Db 454 GlnThrProgluInrSerProAspIySgluGluAlaSerIleProIySlySleuGlnThr 473

QY 1426 GCCAGGCTTGAAGACAGCAGAGCAAGGATTCATTCGTCAGAACAGCAGCAGTGGC 1485
   |||||
Db 474 AlaArgProSerAspSerGlnSerIySArgleuHisSerValArgThrSerSerIy 493

QY 1486 ACCGCCAGAGGTCCTTTTATCTTCACATGCATCGAAGCTGGAGCGGAGACAATTAC 1545
   |||||
Db 494 ThrAlaGlnArgSerleuIeuserProleuHisArgSerGlySerValGluuAspInrYr 513

QY 1546 CACACACACTTCTTTTGGGCTTTCACACAGCCAGCAGACACCTCAGAAAGCTGTCGC 1605
   |||||
Db 514 HisThrIserPheleuPheGlyIeuserThIrgInglnHisIleuThrIySserAlaGly 533

QY 1606 CTGGGCTTTAAGGGGTGGCAGCTCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTG 1665
   |||||
Db 534 IeuglyIeuySglYrIrgHisSerAspIleuAlaProGlnInrThrIProSerIe 553

QY 1666 ACCAGACAGTGTATTTTGCACAGAGTCTCACAATTCTACTGCTCAGCCATCTAC 1725
   |||||
Db 554 ThrIserIerIyIrgPheAlaIthrGluIeuserHisIySerIySerAlaIleYr 573

QY 1726 GAGGCACTGCCAGTACTCTGCCCTCAAGCTGCAGCCAGCTGCCCACTTGGGAGACCA 1785
   |||||
Db 574 GlyIySserAlaSerIySerAlaIyIserCysSerGlnIuProThrCysGlyAspGln 593

QY 1786 GTCATTCGTGGCGAGCGGAGAGCAAGTGAAGAGTGAAGAGTGAAGTGCAGGAGCTGG 1845
   |||||
Db 594 ValIySerValArgArgIrgInIySProSerAspAlaIaAspSerArgIrgSerTrp 613

QY 1846 CATGAAGAGACCCCTTTGAAAAGCAGTTTAAACGACAGACCTGCCAAATGGAATTTGGA 1905
   |||||
Db 614 HisGluGluIeuserProPheGluYlYleGlnPheYsArgIrgSerCysGlnMetGluPheGly 633

QY 1906 GAGACATCATGTCAAGAAACAGTCAAGGAGAGAGAGCTGGGAGAAAGTGGAGTCACT 1965
   |||||
Db 634 GuIserIleMetSerGluuAsnArgSerArgIrgIuIeuglyIySvalGlyIserGlnSer 653

QY 1966 AGCTTTCGGGAGCATGAATCATTAGGCTTCC 2001
   |||||
Db 654 SerPheSerGlySerMetGlnIleIleGluValIser 665

RESULT 10
US-10-257-026-2
; Sequence 2, Application us/10257026
; Publication No. US20040086659A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10XDMs
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
```


TYPE: PRT
ORGANISM: Homo sapiens
US-10-029-345a-108

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0	665	472	472	0	0	0
Score:	472.00	100.00%	100.00%	0	0	0
Best Local Similarity:	100.00%	100.00%	100.00%	0	0	0
Query Match:	70.76%	16	0	0	0	0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-10-257-026-2 (1-665)

```

QY 586 ACCGTGCCAAAGCTGATCTTATCCCGAGTCTCATTTCTGGGTGGCTGTGAATGAC 645
DB 194 ThnCysProlyserProbaPheilleProGluSerHlePheLeuArgValProValaAsp 213
QY 646 AGCTTTGTGAGAAATTTTGGCGTGTGAGCAAAATCAGTAGATTTCATTGAGAAAGCA 705
DB 214 SerPheCysGluIuylleuProTPrleuAspIysSerValaAspPheilleGluIuylsAla 233
QY 706 AAAGCTCCAAAGATGTGTTCAGTGTGACCTGTTAGCTGGAATCTCCGCGCCGAC 765
DB 234 LysAlaSerAengIyCysValleuValHleCysleuHlaGlylleSerArGserAlaThr 253
QY 766 ATGGCTATGGCTTCATCATGAAAGAGATGACATGCTTTAGTGAAGCTTACAGATT 825
DB 254 lAlAllellelAtyrlleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 273
QY 826 GTGAAGAAAGAAAGCTATATATCTCAAACTTCAATTTTCCGGGCAACTCTGGAC 885
DB 274 ValIleGluIuylsArgProHlePheProAspPheAspPheIleuIleuIleuIleuIleu 293
QY 886 TATGAGAAAGATTAAGAACAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTG 945
DB 294 TyrgIuylsArgIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 313
QY 946 CTGCACCTGGAGAAAGCAATGAATGATCTGCTGCTGTCTCAAGGGTGGACGAAAGC 1005
DB 314 leuHleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 333
QY 1006 GAGAGCGCCCTCAGTCCAGCTGCGGCACTGTCTCAATCTCAGAGGAGGAGCAAAAG 1065
DB 334 GluThrProleuSerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArg 353
QY 1066 CCGGTGATCCCGCAGCGCTGCCAGCGTCCGAGCTGAGCGCGTGTGAGAGGAC 1125
DB 354 ProValHlePheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 373
QY 1126 AGCCCGCTGTACAGCGCTCAGTGGCTGACCTGTCCGACAGACAGAGCTGAAAGACAGC 1185
DB 374 SerProleuValAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 393
QY 1186 AATAAGGTCAGAGGCTCTCTCTCTGATATCAATCAATTCATTTAGAGGACAGAT 1245
DB 394 AsnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 413
QY 1246 GCAGCATCTTACATGGCTTCTCTCTCATCAGAGATGCTTTGAAATTAATAACCTTCC 1305
DB 414 AlaAlaSerleuHleGlyPheSerSerSerSerSerSerSerSerSerSerSerSerSerSer 433
QY 1306 ACTACTCTGGATGGAGCAACAAAGCTATGCGAGTTCCTCCCTGTTCAAGAACTATCGAG 1365
DB 434 ThrThrleuAspGlyThrIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 453
QY 1366 CAGACTCCGAAACCGACTCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1425
DB 454 GluThrProGluThrSerProbaPheilleProGluSerHlePheLeuArgValProValaAsp 473
QY 1426 GCCAGGCTTCAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1485
DB 474 AlaArgProSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 493

```

```

QY 1486 ACCGCCAGAGGACCTTTTATCTCAGTGCATGAAAGTGGAGCGGAGAGCAAAATAC 1545
DB 494 ThrAlaGluThrSerleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 513
QY 1546 CACACAGGCTTCCTTTTGGCGCTTTCCACAGGACAGACAGACCTTCAAGATCTGTGC 1605
DB 514 HleThrSerPheleuHleGlyIleuSerThrSerGlnHleIleuThrIleuSerAlaGly 533
QY 1606 CTGGGCTTAAAGGCTGGCACTCGATATCTTGGCCCCAGACCTTACCTTCCCTG 1665
DB 534 leuGlyleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 553
QY 1666 ACCAGAGCTGGATTTTCCACAGAGCTCTCAGACTTCTACCTGCTCCAGCATCTAC 1725
DB 554 ThrSerSerTrpThrPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 573
QY 1726 GGAGCAGTGGCAGTACTCTGCTCAGCTGACCTGAGCAGCTGCCACTTGGCAGAGCA 1785
DB 574 GlyGlySerAlaSerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 593
QY 1786 GTCTATCTGTGGCGAGCGCGAGAGCAAGTACAGAGCTGACCTGCGCGGAGCTGG 1845
DB 594 ValIleSerValArgArgArgGlnIleuProSerAspArgAlaAspSerArgArgSerTrp 613
QY 1846 CATGAAGAGAGCCCTTTGAAAGAGCTTTAAACGAGAGAGCTGCCAAATGGAATTGGA 1905
DB 614 HleGluIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 633
QY 1906 GAGAGCATCATGTCAGAGAAAGAGTCAAGAGAGAGCTGGAGAGAGAGAGAGAGAGAG 1965
DB 634 GluSerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 653
QY 1966 AGCTTTGGCGAGCATGGAATCATGAGTCTCC 2001
DB 654 SerPheSerGlySerMetGluIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 665

```

RESULT 11

US-10-648-593-240
Sequence 240, Application US/10648593
Publication No. US20040106132A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
FILE REFERENCE: D0273 NP
CURRENT APPLICATION NUMBER: US/10/648,593
PRIOR FILING DATE: 2003-08-26
PRIOR FILING DATE: 2002-08-27
NUMBER OF SEQ ID NOS: 557
SOFTWARE: PatentIn version 3.2
SEQ ID NO 240
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-10-648-593-240

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0	665	472	472	0	0	0
Score:	472.00	100.00%	100.00%	0	0	0
Best Local Similarity:	100.00%	100.00%	100.00%	0	0	0
Query Match:	70.76%	16	0	0	0	0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-10-648-593-240 (1-665)

```

QY 586 ACCGTGCCAAAGCTGATCTTATCCCGAGTCTCATTTCTGGGTGGCTGTGAATGAC 645
DB 194 ThnCysProlyserProbaPheilleProGluSerHlePheLeuArgValProValaAsp 213

```


Db 294 TyrglulysyllylelysaenglnthrglylaserglyProlyserlyleuylleu 313
Qy 946 CTGCACCTGGAGAGCCAAATGAACTGTCCCTGTCTCTCACAGGGTGGACGAAAGC 1005
Db 314 LeuhlsleuyllylsProaengluProvalProalvalasergluglyglnlysser 333
Qy 1006 GAGAGCCCTTCAGTCCAGCTGTGCGCACTGTCTCACTCAGAGGAGAGAGCAAG 1065
Db 334 GluthrProleuserProProcybAlaapSerlathrsergluaalagllyglnly 353
Qy 1066 CCCGTGCATCCCGCAGCGTCCAGCGTCCAGCGTCCAGCGTCCGTGTTAGAGAC 1125
Db 354 ProvalHisProalaserValProserValProserValGlnProserleuGluuap 373
Qy 1126 AGCCCGTGTGTACAGGGCTCAGTGGCTCAGCTGTCCCGACAGAGGCTGGAAAGAC 1185
Db 374 SerProleuValGlnAlaLeuSerGlyLeuHlsleuSerlAlaapArgleuGluuapSer 393
Qy 1186 AATAGCTCAAGCGTTCCTCTCTGTGATATCAATCAGTTTCATATTCAGCCAGCATG 1245
Db 394 AsnlyleuylsArtsrProserleuAspIlelyserValIsertyserAlaSerMet 413
Qy 1246 GCAGCATCTTACATGGCTTCTCTCTCATAGAAAGATCTTTGGAATACTAACCTTCC 1305
Db 414 AlaAlaserleuHlsleuGlyPheSerSerSerGlnuapAlaLeuGlnlyrlyrlyProser 433
Qy 1306 ACTACTGTGATGGAGCAACAAAGCTATGCGCAGTTCCTCCCTGTCCAGAACTATGCGAG 1365
Db 434 ThrThrleuAspGlyThrAsnlyleuGlyAsglPheSerProvalGlnGluuSerGlu 453
Qy 1366 CAGACTCCGAAACCACTCTGTATAGAGAGAGCCAGCTCCCGCAAGAGCTGCAGACC 1425
Db 454 GlnThrProgluthrSerProserAspGlyGlnGluAlaserIleProlysllyleuGlnThr 473
Qy 1426 GCCAGGCTTCAGACAGCCAGAGCAAGCATGTCATTCGTCAGAACACAGCAGCATGCGC 1485
Db 474 AlaArgProserAspSerGlnserlyArglleuHlsSerValArglThrserSerSerGly 493
Qy 1486 ACCGCCAGAGGTCCTTTTATCTCCATGTCATGGAAGTGGAGCGTGGAGCAATTAC 1545
Db 494 ThrAlaIlnArgerleuLeuSerProleuHlsArgerlySerValGluuAspAsnlyr 513
Qy 1546 CACACGACTTCCTTTTGGGCTTTCCACAGCAGAGAGCACTCCAGAGTTCGCTGCGC 1605
Db 514 HisThrSerPheleuPheGlyleuSerThrSerGlnlnHlsleuThrlysserAlaGly 533
Qy 1606 CTGGGCTTTAAGGCTGGCACTCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTG 1665
Db 534 LeuGlyleuylsGlyTrpHisSerAspIleleuAlaProGlnThrserThrProserleu 553
Qy 1666 ACCAGCAGCTGTGATTTTGGCAAGAGTCTTCACACTTCTACTCTGCTCAGCCATCTAC 1725
Db 554 ThrSerSerTrpTyrrPheAlaThrGlnSerSerHisPheTyrrSerAlaSerAlaIleTy 573
Qy 1726 GAGAGCGAGCGCACTTCTCGCTCAGAGTGCAGCGAGTGGCCACTTGGAGAGCCAA 1785
Db 574 GlyGlySerAlaSerlyrSerAlaTyrrSerCysSerlInuProThrCysGlyAspGln 593
Qy 1786 GTCATATCTGTGGCAGAGCGGAGAGCCAAAGTACAGAGCTGACCTGCGCGGAGCTGG 1845
Db 594 ValTyrrSerValArArArGlnlyrProserAspArgAlaapSerArArArGlnlyr 613
Qy 1846 CATGAGAGAGCCCTTTGAAAAGAGCTTTAAAGCAGAAAGCTGCCAAATGGAATTTGA 1905
Db 614 HisGlnIuSerPerProPheGlnlyrGlnPhelelyrArgrSerCysGlnMetGluPheGly 633
Qy 1906 GAGAGCATCATGTCAGAGAGAGTCAAGGAGAGAGTGGGAGAAAGTGGCACTGCTC 1965
Db 634 GluSerlleuSerSerGluuAsnArgrSerArGlnGlnlyleuGlylyValGlySerGlnSer 653
Qy 1966 AGCTTTTGGGAGCATGAAATCATTTAGAGTCTCC 2001
Db 654 SerPheSerGlySerMetGlnIlelleGlnValSer 665

RESULT 13
US-10-072-012-679
Sequence 679, Application US/10072012
Publication No. US2004003493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernhusen, Bryan
APPLICANT: Patursajan, Meera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Beha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimír Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 679
LENGTH: 690
TYPE: PR1
ORGANISM: Homo sapiens
US-10-072-012-679
Alignment Scores:
Pred. No.: 0
Score: 472.00
Percent Similarity: 100.00%
Beet Local Similarity: 100.00%
Query Match: 70.76%
DB: 12
Length: 690
Matches: 472
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0
US-10-029-345a-108_COPY_532_2532 (1-2001) x US-10-072-012-679 (1-690)
Qy 586 ACCTGTCCAAAGCTGACTTATCCCGAGTCTCATTTCTGTGCTGTGATGAC 645

Db 219 ThrCysProLysProAspPheIleProGluSerHisPheLeuArgValProValAsnAsp 238
Qy 646 AGCTTTTGAGAAATTTTGGCGTGGTGGACAAATCAGTAATTTTCATTGGAGAAAGCA 705
Db 239 SerPheCysGluLysIleLeuProTyrPheLysAspLysSerValAspPheIleGluLysAla 258
Qy 706 AAAGCCTCAATGAGATGTCTTCTAGTGCATGTTTGGCTGGATCTCCCGCTCCGACAC 765
Db 259 LysAlaSerAsnGluCysValLeuValHisCysLeuAlaGluLysLeuSerArgAlaThr 278
Qy 766 ATGCTATGCCCTTACATCATGAAGAGATGGAACATGCTTTTAAATGAAGCTTACAGATT 825
Db 279 ILeAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPhe 298
Qy 826 GTGAAGAAAAAGACCTATATCTCCAACTTCAATTTTGGGCGCAACCTCCGAGAC 885
Db 299 ValLysGluLysArgProIleHisSerProAsnPheAsnPheLeuGluGlnLeuLeuAsp 318
Qy 886 TATGAGAAGAATTAAGAACAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTG 945
Db 319 TyrGluLysLysIleLysAsnGlnThrGluLysSerGlyProLysSerLysLeuLysLeu 338
Qy 946 CTGCACCTGGAGAAAGCCAAATGAACCTGTCTCTGCTGTCTCAGAGGCTGACAGAAAGC 1005
Db 339 LeuHisLeuGluLysProAsnGluProValProAlaValSerGluGluGlnLysSer 358
Qy 1006 GAGAGCGCCCTGACGACCCGTGGCCGACCTGTGCTACCTCAGAGGCGAGAGCAAAAG 1065
Db 359 GlnThrProLeuSerProProCysAlaAspSerAlaThrSerGlnAlaGluGlnArg 378
Qy 1066 CCGGTGCATCCCGCAGCGTCCAGCGTGCACGAGCGTGACGCTGCTTGAAGAGAC 1125
Db 379 ProValHisProAlaSerValProSerValProSerValGlnProSerLeuLeuGluAsp 398
Qy 1126 AGCCCGCTGGTAAAGCGCTCACTGAGGCTGACCTGTCCGACAGACAGCTGAAAGACAC 1185
Db 399 SerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSer 418
Qy 1186 AATAAGCTCAAGGCTCTCTCTCTGAGTATCAATCAATTCAGTTCAGCCGACAG 1245
Db 419 AsnLysLeuLysArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMet 438
Qy 1246 GCAGCATCTTACATGAGCTTCTCTCATGAGAAAGATGTTGGAACTACTACAACTTCC 1305
Db 439 AlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGluLysTyrTyrLysProSer 458
Qy 1306 ACTACTCTGATGGAGCAACAAAGCTATGCAAGTTCTCCCTGTTGAGAACTATTCGAG 1365
Db 459 ThrThrLeuAspGluLysHisLysLeuGluGlnPheSerProValGlnGluLeuSerGlu 478
Qy 1366 CAGACTCCCGAAACAGAGCTCTGATTAAGAGAAAGCAGAGATCCCAAGAAAGCTGACAG 1425
Db 479 GlnThrProGluLysSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThr 498
Qy 1426 GCGAGGCTTACAGACAGCAGAGCAAGCATTTGCAATTCGATGGAACAGCAGAGCTGAC 1485
Db 499 AlaArgProSerAspSerGlnSerLysArgLeuHisSerValArgHisSerSerSerGly 518
Qy 1486 ACCGCCCAAGATCCCTTTTATCTCACTGCATCGAAGTGGAGCGGTGAGACAATTAC 1545
Db 519 ThrAlaGlnArgSerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyr 538
Qy 1546 CACACAGATCTCTTTGGGCTTTCCACAGCCAGACACACTCAACAAAGCTGTCTGAC 1605
Db 539 HisThrSerPheLeuPheGluLysLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGly 558
Qy 1606 CTGAGGCTTAAAGGCTGAGCTCGATATCTTGGCCCGCAGACCTTACCCCTTCCCG 1665
Db 559 LeuGluLysLeuLysGlyTyrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeu 578
Qy 1666 ACCAGAGCTGATTTTCCACAAGAGTCTTCACTTACTGCTCAGCCACTTAC 1725

Db 579 ThrSerSerTyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyr 598
Qy 1726 GAGAGCAGTGCAGATTAATCTTCTGCTTACAGCTCCAGCCAGCTCCGACATTCGAGAACCA 1785
Db 599 GlyGlySerAlaSerTyrSerAlaTyrSerLysSerGlnLeuProThrCysGluLysPgl 618
Qy 1786 GTCTATTTCTGTGCGAGCGCGCAGAGCAAGTGAACAAGCTGCACTCCGCGCGAGCTGG 1845
Db 619 ValTyrSerValArgArgArgGlnLysProSerArgAlaAspSerArgArgSerTyr 638
Qy 1846 CATGAAGAGAGCCCTTTGAAAAGCATTTAAACGACAGAAAGCTGCCAAATGGAATTTGCA 1905
Db 639 HisGluLysSerProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGly 658
Qy 1906 GAGACATCAGTGCAGAAACAGTGCAGGGAAGAGCTGGGGAAGTGGGAGTCACT 1965
Db 659 GluSerIleMetSerGluAsnArgSerArgGluGluLysValGlySerGlnSer 678
Qy 1966 AGCTTTTGGGCGAGCATGGAATCATTTGAGGCTTCC 2001
Db 679 SerPheSerGlySerMetCgluIleIleGluValSer 690

RESULT 14
US-10-072-012-703
; Sequence 703, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchervet, Vellizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patuturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taubier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Groesse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Daniel M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine B.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459

```
/ PRIOR FILING DATE: 2001-02-08
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1391
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 703
/ LENGTH: 690
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-072-012-703

Alignment Scores:
Pred. No.: 0          Length: 690
Score: 472.00        Matches: 472
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 70.76%          Indels: 0
DB: 12                  Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-10-072-012-703 (1-690)

QY 586 ACCGTGTCMAAGCTGACTTATCCCGAGCTCATATTCTGCGTGTGCTGGAATGAC 645
DB 219 ThCybProlysrProspRheIIeProGluSerHisPheLeuArgValProValaAsnAsp 238
QY 646 AGCTTTGTGAGAAATTTTGGCGGTGGACAAATCAGTAGATTTCATTGAGAAAGCA 705
DB 239 SerPheCysGluLysIleLeuProTrpLeuAspLysSerValaAspHeIIeGluLysAla 258
QY 706 AAAGCTCCAAATGATGTGTTTCTAGTGCACCTGTTTGAAGTGTCCGCTCCGCGCAC 765
DB 259 LysAlaSerAsnGluCysValLeuValHisCysLeuAlaGlyLLeSerArgSerAlaThr 278
QY 766 ATGCTATGCTTACATCATGAAGAGATGACATGTTTAAATGAAGTTTCAAGATT 825
DB 279 IleAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPhe 298
QY 826 GTGAAGAAAGAAAGACTTACTATATCTCCAAACTTCATTTTGGGGCCAACTCTGAGC 885
DB 299 ValLysGluLysArgProThrIleSerProAsnPheAsnPheLeuGluInLeuLeuAsp 318
QY 886 TATGAGAAGAGATTAAGAACAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTG 945
DB 319 TyrGluLysLysIleLysAsnGluInThrGluAlaSerGlyProLysSerLysLeuLys 338
QY 946 CTCGACCTGGAGAAAGCAAAATGAATGACTGTCTGCTGCTCTCAAGGGTGGACGAAAGC 1005
DB 339 LeuHisLeuGluLysArgProAsnGluProValProAlaValSerGluGlyInLysSer 358
QY 1006 GAGAGCGCCCTGAGTCCAGCTGGCGGACTGTGTAACCTCAGAGGAGGAGCAAAAG 1065
DB 359 GluThrProLeuSerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArg 378
QY 1066 CCCGTCATCCCGCCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGC 1125
DB 379 ProValHisProAlaSerValaProSerValaProSerValaGlnProSerLeuLeuGluAsp 398
QY 1126 AGCCCGCTGTATACAGCGCTCAGTGGCTCAGCTGTCCGACAGACAGCTGGAAGCAGC 1185
DB 399 SerProLeuValaGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSer 418
QY 1186 AATTAAGCTCAAGGGTCTCTCTCTGATATCAATTAATCAATTCATATTAAGCAGATG 1245
DB 419 AsnLysLeuLysArgSerPheSerLeuAspIleLysSerValaSerTyrSerAlaSerMet 438
QY 1246 GCAGACCTCTTACATGAGCTTCTCTCATCAGAGATGCTTTGGAATTAATCAAACTTCC 1305
DB 439 AlaAlaSerLeuHisGlyPheSerSerSerSerGluAspAlaLeuGluLysTyrLysProSer 458
QY 1306 ACTACTCTGGATGGAGCAACAAGCTATGCCAGTTCTCCCTGTTTCAAGAACTATGGAG 1365
DB 459 ThrThrLeuAspGlyThrIleAsnLysLeuCysGluPheSerProValGlnGluLeuSerGlu 478
QY 1366 CAAAGCTCCGAAACCAAGTCTGATTAAGAGAGAAAGCAAGCTCCCAAGAAAGCTGCAAGC 1425
```

```
DB 479 GluThrProGluThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThr 498
QY 1426 GCCAGGCTTTCAGACAGCCAGAGCAAGCATTTGCATTCGCTGACGAACCAAGCACTAGC 1485
DB 499 AlaArgProSerAspSerGlnSerLysArgLeuHisSerValaArgThrSerSerGly 518
QY 1486 ACCGCCCAGAGGTCCTTTTATCTCCATGCACTGCAAGTGGAGGCTGGAGACAAATTAC 1545
DB 519 ThrAlaGlnArgSerLeuLeuSerProLeuHisArgSerGlySerValaGluAspAsnTyr 538
QY 1546 CACACCAAGCTTCTTTTCGACCTTTTCACAGCCAGCAGCAGCAGCTCAGAACTGTGTGC 1605
DB 539 HisThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGly 558
QY 1606 CTGGGCTTAAAGGCTGGGACCTCGCATTCCTGGCCCCCAGACCTTACCTTCCCTG 1665
DB 559 LeuGlyLeuLysGlyTyrPheSerAspIleLeuAlaProGlnThrSerThrProSerLeu 578
QY 1666 ACCAGAGCTGGATTTTGGCAGAGTCTCAGACTTCTACTGCTGAGCATCTAC 1725
DB 579 ThrSerSerTyrPheAlaThrGluSerSerHisAspHerySerAlaSerAlaIleTyr 598
QY 1726 GAGAGCAGTGCAGTACTCTGCTCCTCAGCTGACAGCCAGCTGCCACTTGGCGAGACAA 1785
DB 599 GlyGlySerAlaSerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGln 618
QY 1786 GTCTATTCTGTGGCGAGCGCGCAAGAACAGTACAGACTGACTCGCGCGGAGCTGG 1845
DB 619 ValTyrSerValaArgArgArgGlnLysProSerAspArgAlaAspSerArgAspSerTyr 638
QY 1846 CATGAAGAGAGCCCTTTTAAAGCAAGTTTAAACGAGAAAGCTGCCAAATGGAATTTGGA 1905
DB 639 HisGluGlnSerProGluGluGlnPheLysArgArgSerCysGlnMetGluPheGly 658
QY 1906 GAGAGCATATGTCAGAGAAACAGTTCACCGGAAGAGCTGGGAAAGTGGGCACTAGCT 1965
DB 659 GluSerIleMetSerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSer 678
QY 1966 AGCTTTGCGGCGAGCATGGAAATCATTAAGAGTCTCC 2001
DB 679 SerPheSerGlySerMetGluIleIleGluValSer 690

RESULT 15
US-10-425-114-54204
/ Sequence 54204, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaka, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT FILING DATE: US/10/425,114
/ PRIOR FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 54204
/ LENGTH: 690
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4119-028-H6_F01 pep
US-10-425-114-54204

Alignment Scores:
Pred. No.: 0          Length: 690
Score: 472.00        Matches: 472
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
```

Query Match: 70.76% Indels: 0
DB: 12 Gaps: 0
US-10-029-345a-108_COPY_532_2532 (1-2001) x US-10-425-114-54204 (1-690)

QY 586 ACTGTCCAAAGCTGACTTATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGAC 645
DB 219 ThrcysProIysProAspPheIleProGluSerHisPheLeuArgValProValAsnAsp 238
QY 646 AGCTTTGTGAGAAATTTTGGCTGTGGACAAATCAGTAAGATTTCATTGAGAAAGCA 705
DB 239 SerPheCysGluIysIleLeuProTyrLeuAspIysSerValAspPheIleGluIysAla 258
QY 706 AAAGCCTCAATGGATGTGTTCTAGTCACATGTTTACCTGGAGATCTCCGCTCCGACACC 765
DB 259 LysAlaSerHisnGlyCysValIleLeuValHisCysLeuAlaIleSerHisSerAlaThr 278
QY 766 ATGCTATCGCCTCATCATGAGAGAGATGACATGCTTTAGATGAGCTTACAGATT 825
DB 279 ILeAlaIleAlaTyrIleMetIysArgMetAspMetSerLeuAspGluAlaTyrArgPhe 298
QY 826 GTGAAGAAAGAAAGCCTTATATCTCCAACTTCATTTTCTGGGCCAACTCTTGAC 885
DB 299 ValIysGluIysArgProIthrIleSerProAsnPheAsnPheLeuGluIleLeuAsp 318
QY 886 TATGAGAGAGAGATTAAGAACAGACTGAGCATCAGGGCCAAAGAGCAATCAAGCTG 945
DB 319 TyrGluIysIysIleIysAsnGlnThrGlyAlaSerGlyProIysSerIysLeuIysLeu 338
QY 946 CTGCACTGTGAGAGAGCCAAATGAACCTGTCTCTGCTGTCTCAGAGGATGAGAGAAAGC 1005
DB 339 LeuHisIleuGluIysArgProAsnGluProValProAlaValSerGluIysGlnIysSer 358
QY 1006 GAGAGCGCCCTGAGTCACCGCTGTGCGCACTGTGCTACCTCAGAGGACAGAGCAAGG 1065
DB 359 GluThrProIysSerProIocysAlaAspSerAlaThrSerIuAlaIleGlyIlnArg 378
QY 1066 CCCGTGCATCCCGCAGCGTCCCAAGCAGTGCAGCGTGCAGCGCTGCTTGAAGAGAC 1125
DB 379 ProValHisProAlaSerValProSerValProSerValGlnProSerLeuLeuIuAsp 398
QY 1126 AGCCCGCTGTGACAGCGCTCATGTGGCTGTGACCTTCCGACAGACAGGCTGAGAGAC 1185
DB 399 SerProLeuValGlnAlaLeuSerGlyLeuHisIleLeuSerAlaAspArgLeuGluAspSer 418
QY 1186 AATAAGCTCAAGGCTTCTCTCTGTGATATCAATTCATTTCATATTCACGCCAGCAG 1245
DB 419 AsnIysLeuIysArgSerPheSerIleuAspIleIysSerValSerTyrSerAlaSerMet 438
QY 1246 GCAGCATCTTATCATGCTTCTCTCATGAGAGATGCTTTGGAATATCAAACTTCC 1305
DB 439 AlaAlaSerLeuHisIleIysPheSerSerSerGluAspAlaIleuGluTyrTyrIysProSer 458
QY 1306 ACTACTCTGTGATGGACCAACAAGCTATAGCCAGTTTCCCTGTTCAAGAACTATGAG 1365
DB 459 ThrThrLeuAspGlyThrAsnIysLeuCysGlnPheSerProValGlnIuIuLeuSerGlu 478
QY 1366 CAGACTCCCGAAACAGCTCTGATAGAGAGAGAGCAGATCCCAAGAAGCTGAGAGCC 1425
DB 479 GlnThrProGluThrSerProAspIysGluGluAlaSerIleProIysIysLeuGlnThr 498
QY 1426 GCAGAGCCTTCAAGACAGCAGAGCAAGCATTCATTCGTCAGAACAGCAGCAGTGGC 1485
DB 499 AlaArgProSerArgSerIleSerIysArgLeuHisIleSerValArgThrSerSerIys 518
QY 1486 ACCGCCAGAGGTCCTTTATCTTCACATCGATCCAAAGTGGAGAGGTGAGAGCAATTAC 1545
DB 519 ThrAlaGlnArgSerIleuLeuSerProLeuHisAspSerGlySerValGlnAspAsnTyr 538
QY 1546 CACACAGCTTCTTTGGCTTTCACACAGCAGCAGCAGCAGCTTCAAGAAAGCTGCTGGC 1605
DB 539 HisThrIserPheLeuPheGlyLeuSerThrIserGlnHisIleLeuThrIysSerAlaGly 558

QY 1606 CTGGGCTTAAGGCTGGACACTCGATATCTTTGGCCCCCAGACCTTACCCCTTCCCTG 1665
DB 559 LeuGlyLeuIysGlyTyrPheIleuAlaIleuAlaProGlnThrSerThrProSerLeu 578
QY 1666 ACCAGAGCTGTGATTTTGGCCACAGAGTCTCAGACTTCTACTGCTGACCCATCTAC 1725
DB 579 ThrSerSerTyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyr 598
QY 1726 GAGAGCAGTGGCAGTTACTCTGCTTACAGCTGACCCAGCTGCCCATTTGCCGAGACCA 1785
DB 599 GlyIysSerAlaSerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGln 618
QY 1786 GTCATATTCGTGGAGGCGGAGAGCCAGATGACAGAGCTTCCGCGGCGAGCTGG 1845
DB 619 ValTyrSerValArgArgArgGlnIysProSerAspArgAlaAspSerArgArgSerTyr 638
QY 1846 CATGAGAGAGCCCTTTGAAAGAGATTAAACCCAGAAAGCTGCCAAATGAAATTTGGA 1905
DB 639 HisGluGluSerProPheGluIysGlnPheIysArgArgSerCysGlnMetGluPheGly 658
QY 1906 GAGAGCATCATGTCAGAGAACAGTCAAGGAGAGCTGGGGAAAGTGGGCAAGTCACT 1965
DB 659 GluSerIleMetSerGluAsnArgSerArgGluGluLeuGlyValGlySerGlnSer 678
QY 1966 AGCTTTGGGCGAGCATGGAATCATTTGAGGCTTCC 2001
DB 679 SerPheSerGlySerMetGluIleIleGluValSer 690

RESULT 16
US-10-072-012-258
Sequence 258, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernov, Vellizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernusen, Bryan
APPLICANT: Patuzajan, Meera
APPLICANT: Shinkens, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esna
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr., Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosee, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767


```

/ APPLICANT: Alsobrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072,012
/ PRIOR FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 60/266,767
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: 60/267,057
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/266,975
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/267,459
/ PRIOR FILING DATE: 2001-02-08
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1391
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 256
/ LENGTH: 680
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-072-012-256

Alignment Scores:
Pred. No.: 0 Length: 680
Score: 394.00 Matches: 394
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.07% Indels: 0
DB: 12 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-10-072-012-256 (1-680)
QY 820 AGATTGTGAAAGAAAAAGACCTACATATCTCAAACTTCAATTTTCTGGGCCCAACTC 879
DB 287 ArgPheValIysGluYsaIgrProthIleSerProasnPhenasnPhenLeuGIYglIneu 306
QY 880 CTGACATATGAGAAAGAAAGATTAAAGAACAGATGAGCATCAAGGCCCAAGAAAGAACTC 939
DB 307 LeuAspTyrGIYglYsIleYsaInGIthGIYAlaSerGIYProIysSerIYsIleu 326
QY 940 AAGCTGTGACCTGGAGAAAGCCAAATGAACCTGCTCCCTGCTGTCTCAGAGGCTGACAG 999
DB 327 LysLeuLeuHISleuGIYsIleYsaInGIuProValProIAlaValSerGIYglYglIn 346
QY 1000 AAAAGGAGAGCGCCCTCACTGACCTGACCTGCGCACTGTGTAACCTCAGAGGACAGAGA 1059
DB 347 LysSerGIuThrProIeuserProProCYsaIaAspSerAlaThSerGIuAlaIaGIY 366
QY 1060 CAAAGCCCGGTGATCCCGCCAGCGCTGCCAGCGTGCAGAGGTGACAGCGCTGCTTTA 1119
DB 367 GlnArgProValHisProIaSerValIProSerValIProSerValGIuProSerIeuser 386
QY 1120 GAGAGACGCCCGGTGTATAGAGCGCTCAGTGGCTGCAGCACTCTCCGAGACAGAGCTGAA 1179
DB 387 GluAspSerProIeuserValGIuAlaIeuserGIYleuHISleuSerValaAspArgIleuGIu 406
QY 1180 GACAGCAATTAAGCTCAAGCGTCTCTCTCTGATATCAATCAGTTTCATTTATTCAGCC 1239
DB 1180 GACAGCAATTAAGCTCAAGCGTCTCTCTCTCTGATATCAATCAGTTTCATTTATTCAGCC 1239
```

```

DB 407 AspSerAsnIYsIleuYsaIgrSerPheSerIleuAspIleYsSerValSerIYsSerAla 426
QY 1240 AGCATGGACAGCATCTTATACATGAGCTTCTCTCATTCAGAAAGATGCTTGAATACACAA 1299
DB 427 SerMetAlaIaSerIeuserIleGIYsPheSerSerSerGIuAspIaIeuserIuYIYrIYs 446
QY 1300 CTTTCCACATCTCTGATGGAGCAACAAAGCATATGCACTTCTCCCTGTTGAGAACTA 1359
DB 447 ProSerThrThrIeuserAspGIYthrasnIYsIleuCYsGIuInPheSerProValGIuInIeuser 466
QY 1360 TCGAGCAGACCTCCGAAACCAAGTCCTGATAGAGAGAAAGCCAGCATCCCAAGAACTG 1419
DB 467 SerGIuInThrProGIuThrSerProAspIYsGIuAlaIaSerIleProIYsIYsIeuser 486
QY 1420 CAGACCGCAGGCGCTTACAGACAGCAGCAAGCATTCGATTCGATCGAGAAACGACAGC 1479
DB 487 GlnThrAlaArgProSerAspSerGIuSerIYsArgIleuHISleuSerValArgThrSerSer 506
QY 1480 AGTGGACCGCCCAAGAGTCCCTTTATCTCACTGCATCGAAGTGGAGCGTGGAGAGC 1539
DB 507 SerGIYthrAlaIaGIuArgSerIeuserIeuserProIeuserIaArgSerGIYSerValGIuAsp 526
QY 1540 AATTACACACAGCAGCTCTTTCGACCTTTCACAGCAGCAGCAGCAGCAGCTTACCAAGACT 1599
DB 527 AsnIYrHISerThSerPheIeuserPheGIYleuSerThSerGIuInHISleuThrIYsSer 546
QY 1600 GCTGGCTGGGCGCTTAAAGGCGTGGAGCTCGAGATCTTGAGCCCGCCAGACCTTACCCCT 1659
DB 547 AlaGIYleuGIYleuYGIYTrpHISerAspIleuAlaProGIuInThrSerThrPro 566
QY 1660 TCCCTGACAGAGAGTGTATTTTGGCCACAGAGCTTCACAGCTTCTACTGCTCTACGCC 1719
DB 567 SerIeuserThrSerThrPheIaThrIeuserThrIaPheIYsSerAlaSerAla 586
QY 1720 ATCTACGAGGACAGTCCAGTACTGCTGCTCAAGCTGACAGCCAGCTGCCCACTTGCGCA 1779
DB 587 IleYrGIYglYsSerAlaSerIYsSerAlaIYsSerIYsSerGIuInProthIYsGIY 606
QY 1780 GACCAAGCTATTTCTGTGCGGAGCGGCGCAAGAGCCAAAGTGAAGAGCTGACTCGGCGG 1839
DB 607 AspGIuValIYsSerValArgArgArgGIuIYsProSerAspArgIaAspSerArgArg 626
QY 1840 AGCTGCGATGAAGAGAGCGCCCTTGAAGAGAGATTTAAAGCGAGAACCTGCCAAATGGA 1899
DB 627 SerTrpHisGIuIeuserProPheGIuIYsGIuInPheYsaIgrYsSerCYsGIuInMetGIu 646
QY 1900 TTGGAAGAGCATCATGTGACAGAAACAGGTCAACGGGAAGAGCTGGGGAAGTGGCGAGT 1959
DB 647 PheGIYgluSerIleuMetSerGIuAsnArgSerArgGIuIeuserGIYsValGIYsSer 666
QY 1960 CAGTCTAGCTTTTGGGCGACATGAGAAATCATTTGAGGTCTCC 2001
DB 667 GlnSerSerPheSerGIYsSerMetGIuIleIleGIuValSer 680

RESULT 18
US-10-296-115-1259
/ Sequence 1259, Application US/10296115
/ Publication No. US20040053248A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq Inc
/ TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides
/ FILE REFERENCE: 784PCT
/ CURRENT APPLICATION NUMBER: US/10/296,115
/ CURRENT FILING DATE: 2002-11-18
/ PRIOR APPLICATION NUMBER: US09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ PRIOR APPLICATION NUMBER: US09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ NUMBER OF SEQ ID NOS: 1478
/ SEQ ID NO 1259
/ LENGTH: 672
/ TYPE: PRT
/ ORGANISM: Homo sapiens
```



```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(672)
; OTHER INFORMATION: Xaa = any amino acid or other as shown in Table 3
US-10-296-115-1259

Alignment Scores:
Pred. No.: 2,55e-238 Length: 672
Score: 258.00 Matches: 258
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38,68% Indels: 0
DB: Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-296-115-1259 (1-672)

QY 586 ACCTGTCGCAAGCTGATTCATTCCTCCGAGCTCATTTCTGCGTGTGCTGTAATGAC 645
DB 201 ThCyProLySerProAlaPheIleProGluSerHisPheLeuArgValProValAsnAsp 220

QY 646 AGCTTTTGAGAAATTTTGGCGTGGTGGACAAATCAGATTCATTGATGAGAAAGCA 705
DB 221 SerPheCysGluLysIleLeuProTyrLeuAspLysSerValAspPheIleGluLysAla 240

QY 706 AAAGCTCCAAATGATGTGTTCTAGTGCACCTGTTAGCTGGATCTCCGCTCCGACACC 765
DB 241 LysAlaSerAsnLysCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThr 260

QY 766 ATGCTATCGCTCATCATCATGAAAGAGATGAGACATGCTTTAGATGAACTTACAGATT 825
DB 261 IleAlaIleAlaTyrIleMetLysArgMetCaspMetSerLeuAspGluAlaTyrArgPhe 280

QY 826 GTGAAGAAAAAAGAACTACTATATCTCCAACTTCAATTTTCTGGGCAACTCTCGAGAC 885
DB 281 ValLysGluLysArgProThrIleSerProAsnPheAsnPheLeuGluLysLeuLeuAsp 300

QY 886 TATGAGAAGAAAGTAAAGAACAGACAGATGAGACATGAGGCGCAAGAACTCAAGCTG 945
DB 301 TyrGluLysLysIleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeu 320

QY 946 CTGACCTGAGAGAGCCAAATGAACTGCTGCTGCTCTCAGAGGGTGGACGAAAGAC 1005
DB 321 LeuHisLeuGluLysProAsnGluProValProAlaValSerGluGlyGlnLysSer 340

QY 1006 GAAAGCGCCCTCAGTCCACCTGTGCGGACATCTGCTTACCTCAAGAGCAGAGCAAAAG 1065
DB 341 GluThrProLeuSerProProCysAlaAspSerAlaThrSerGluAlaIleGlyGlnArg 360

QY 1066 CCGGTGATCCGCGCAGCGTGGCCAGCGTCCAGCGTGGAGCGTGTGTTAGAGGAC 1125
DB 361 ProValHisProAlaSerValProSerValProSerValGlnProSerLeuLeuGluAsp 380

QY 1126 AGCCCGTGTGATCAGGCGCTCAGTGGGCTGACCTGTCCAGACAGCAGCTGAAAGACAG 1185
DB 381 SerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerLysAlaAspArgLeuGluAspSer 400

QY 1186 AATAAGCTCAAGGCTTCTTCTCTGAGATATCAATCAGTTTCATATTCAGCCAGCATG 1245
DB 401 AsnLysLeuLysArgSerPheSerLeuAspIleLysSerValSerLysSerAlaSerMet 420

QY 1246 GCAGCATCTTACATGGCTCTGCTCATCAGAGAGATCTTGAATAATCAATCAAACTTCC 1305
DB 421 AlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGluLysTyrLysProSer 440

QY 1306 ACTACTCTGGATGGAGCAACAAAGCTATGCGCAGTCTTCCCTGTTCAGAACTA 1359
DB 441 ThrThrLeuAspGlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeu 458

RESULT 19
US-10-072-012-682
; Sequence 682, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shimkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Baha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; REMAINING PRIOR APPLICATION DATA REMOVED - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 682
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-072-012-682

Alignment Scores:
Pred. No.: 6,48e-55 Length: 660
Score: 67.00 Matches: 67
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10,04% Indels: 0
DB: Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-072-012-682 (1-660)

QY 49 TTGGTGGCTCTGCTGGAAGTGAACGAAAGTGTCTATGATGACCGGCATTT 108
DB 15 LeuValAlaLeuLeuLysSerGlyThrGluLysValLeuLeuLysSerArgProPhe 34

QY 109 GTGATATCAATATACATCCACATTTTGAAGCCATTATATCAATGCTTCAAGCTTATG 168
DB 35 ValGluTyrAsnThrSerHisIleLeuGluAlaIleAsnIleAsnLysSerLysLeuMet 54

```

QY 169 AAGCGAGTTGCAACAGACAAAGTTAATTACAGAGCTCATCCAGCATTGCGGAAA 228
 |||||
 Db 55 LysArgGluGlnGlnAlaPlysValLeuIleThrGluLeuIleGlnHisSerAlaLys 74
 |||||
 QY 229 CATAAGCTTGACATTGATTGC 249
 |||||
 Db 75 HisLysValAspIleAspCys 81
 |||||
 RESULT 20
 US-10-072-012-683
 ; Sequence 683, Application US/10072012
 ; Publication No. US20040033493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Zerhusen, Bryan
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Li, Li
 ; APPLICANT: Gangolli, Esha
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Anderson, David W.
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Taupier Jr, Raymond J.
 ; APPLICANT: Gusev, Vladimir Y.
 ; APPLICANT: Colman, Steven D.
 ; APPLICANT: Wolenc, Adam R.
 ; APPLICANT: Pena, Carol E. A
 ; APPLICANT: Futak, Katarzyna
 ; APPLICANT: Grose, William M.
 ; APPLICANT: Alsobrook II, John P.
 ; APPLICANT: Lepley, Denise M.
 ; APPLICANT: Rieger, Daniel K.
 ; APPLICANT: Burgess, Catherine E.
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-258
 ; CURRENT APPLICATION NUMBER: US/10/072,012
 ; CURRENT FILING DATE: 2002-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,102
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: 60/265,514
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,517
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,412
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,395
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/266,406
 ; PRIOR FILING DATE: 2001-02-02
 ; PRIOR APPLICATION NUMBER: 60/266,767
 ; PRIOR FILING DATE: 2001-02-05
 ; PRIOR APPLICATION NUMBER: 60/267,057
 ; PRIOR FILING DATE: 2001-02-07
 ; PRIOR APPLICATION NUMBER: 60/266,975
 ; PRIOR FILING DATE: 2001-02-07
 ; PRIOR APPLICATION NUMBER: 60/267,459
 ; PRIOR FILING DATE: 2001-02-08
 ; Remaining Prior Application data removed - See file wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1391
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 683
 ; LENGTH: 677
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-10-072-012-683
 Alignment Scores: 4,91e-52 Length: 677
 Pred. No.: 64.00 Matches: 64
 Score:

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.60% Indels: 0
 Db: 12 Gaps: 0
 US-10-029-345a-108_COPY_532_2532 (1-2001) x US-10-072-012-683 (1-677)
 QY 346 AGTTCACTCTGTTTCACCTGCTTGACAGTGGGCTTGTGAGTTCTTCGTTGTTCCCT 405
 |||||
 Db 114 SerPheAsnSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerArgCysPhePro 133
 |||||
 QY 406 GGCCTCTGTGAAGAAAATCCACTGACCTGACCTGACATTCTTCAGCCTTGTACT 465
 |||||
 Db 134 GlyLeuGlyGlnGlyLysSerThrLeuValProThrCysIleSerGlnProCysLeuPro 153
 |||||
 QY 466 GTTGCACATTGGGCCAACCCGAAATTCCTCCAAATCTTATCTTGCTGCACGAGAT 525
 |||||
 Db 154 ValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyrlLeuGlyCysGlnArgAsp 173
 |||||
 QY 526 GTTCCCAACAG 537
 |||||
 Db 174 ValLeuAsnLys 177
 |||||
 RESULT 21
 US-10-072-012-702
 ; Sequence 702, Application US/10072012
 ; Publication No. US20040033493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Zerhusen, Bryan
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Li, Li
 ; APPLICANT: Gangolli, Esha
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Anderson, David W.
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Taupier Jr, Raymond J.
 ; APPLICANT: Gusev, Vladimir Y.
 ; APPLICANT: Colman, Steven D.
 ; APPLICANT: Wolenc, Adam R.
 ; APPLICANT: Pena, Carol E. A
 ; APPLICANT: Futak, Katarzyna
 ; APPLICANT: Grose, William M.
 ; APPLICANT: Alsobrook II, John P.
 ; APPLICANT: Lepley, Denise M.
 ; APPLICANT: Rieger, Daniel K.
 ; APPLICANT: Burgess, Catherine E.
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-258
 ; CURRENT APPLICATION NUMBER: US/10/072,012
 ; CURRENT FILING DATE: 2002-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,102
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: 60/265,514
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,517
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,412
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,395
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/266,406
 ; PRIOR FILING DATE: 2001-02-02
 ; PRIOR APPLICATION NUMBER: 60/266,767
 ; PRIOR FILING DATE: 2001-02-05
 ; PRIOR APPLICATION NUMBER: 60/267,057
 ; PRIOR FILING DATE: 2001-02-07
 ; PRIOR APPLICATION NUMBER: 60/266,975
 ; PRIOR FILING DATE: 2001-02-07

```

; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 702
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (469)
; OTHER INFORMATION: wherein Xaa is any amino acid.
US-10-072-012-702

Alignment Scores:
Pred. No.: 9.03e-10 Length: 501
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 12 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-072-012-702 (1-501)

QY 730 GTCGACTGTTTACGTGGATCTCCGCTCCGCCACCATGCTATCGCCTACATCATGAG 789
DB 98 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 117

RESULT 22
US-10-072-012-699
; Sequence 699, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
```

```

; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459.
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 699
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-699

Alignment Scores:
Pred. No.: 8.81e-10 Length: 625
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 12 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-072-012-699 (1-625)

QY 730 GTCGACTGTTTACGTGGATCTCCGCTCCGCCACCATGCTATCGCCTACATCATGAG 789
DB 244 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 263

RESULT 23
US-10-072-012-700
; Sequence 700, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
```



```
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57268
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-10-282-122A-57268

Alignment Scores:
Pred. No.: 3 62      Length: 498
Score: 10.00      Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.55%      Indels: 0
DB: 12      Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-282-122A-57268 (1-498)
QY 103 GCGGCTATCAATTAGCAGCACTTTTCCG 74
Db 342 AlAGlYtYcGlIneuAlaAlaLeuPnePro 351

RESULT 26
US-10-225-060-11
; Sequence 11, Application US/10225060
; Publication No. US20030092891A1
; GENERAL INFORMATION:
; APPLICANT: Davis et al.
; TITLE OF INVENTION: Expressed Ligand - Vascular Intercellular Signaling
; FILE REFERENCE: REG 333-Z
; CURRENT APPLICATION NUMBER: US/10/225,060
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US/09/709,188
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 08/740,223
; PRIOR FILING DATE: 1996-10-25
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-060-11

Alignment Scores:
Pred. No.: 301      Length: 503
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.20%      Indels: 0
DB: 14      Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-225-060-11 (1-503)
QY 1083 CGTGCCAGCGTGCCGCGTGCA 1106
Db 80 ArgAlaGlnArgAlaGlnArgAla 87

RESULT 27
US-10-214-812-2
; Sequence 2, Application US/10214812
; Publication No. US20030064470A1
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
; FILE REFERENCE: REG310-K
; CURRENT APPLICATION NUMBER: US/10/214,812
; CURRENT FILING DATE: 2002-08-08
```

```
; PRIOR APPLICATION NUMBER: US/09/202,491
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: PCT/US97/10728
; PRIOR FILING DATE: 1997-06-19
; PRIOR APPLICATION NUMBER: 60/022,999
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: 60/021,087
; PRIOR FILING DATE: 1996-07-02
; PRIOR APPLICATION NUMBER: 08/665,926
; PRIOR FILING DATE: 1996-06-19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-214-812-2

Alignment Scores:
Pred. No.: 300      Length: 509
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.20%      Indels: 0
DB: 12      Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-214-812-2 (1-509)
QY 1083 CGTGCCAGCGTGCCGCGTGCA 1106
Db 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 28
US-10-214-812-3
; Sequence 3, Application US/10214812
; Publication No. US20030064470A1
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
; FILE REFERENCE: REG310-K
; CURRENT APPLICATION NUMBER: US/10/214,812
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: US/09/202,491
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: PCT/US97/10728
; PRIOR FILING DATE: 1997-06-19
; PRIOR APPLICATION NUMBER: 60/022,999
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: 60/021,087
; PRIOR FILING DATE: 1996-07-02
; PRIOR APPLICATION NUMBER: 08/665,926
; PRIOR FILING DATE: 1996-06-19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-214-812-3

Alignment Scores:
Pred. No.: 300      Length: 509
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.20%      Indels: 0
DB: 12      Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-214-812-3 (1-509)
QY 1083 CGTGCCAGCGTGCCGCGTGCA 1106
Db 86 ArgAlaGlnArgAlaGlnArgAla 93
```

```
RESULT 29
US-10-215-224-2
; Sequence 2, Application US/10215224
; Publication No. US20030059887A1
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
; FILE REFERENCE: REG330-K
; CURRENT APPLICATION NUMBER: US/10/215,224
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: PCT/US97/10728
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 60/022,999
; PRIOR FILING DATE: 1997-06-19
; PRIOR APPLICATION NUMBER: 60/021,087
; PRIOR FILING DATE: 1996-07-02
; PRIOR APPLICATION NUMBER: 08/665,926
; PRIOR FILING DATE: 1996-06-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-215-224-2

Alignment Scores:
Pred. No.: 300 Length: 509
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 14 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-215-224-2 (1-509)
QY 1083 CGTCCCGAGCGTGGCCGCGTGCA 1106
Db 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 30
US-10-215-224-3
; Sequence 3, Application US/10215224
; Publication No. US20030059887A1
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
; FILE REFERENCE: REG330-K
; CURRENT APPLICATION NUMBER: US/10/215,224
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: US/09/202,491
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: PCT/US97/10728
; PRIOR FILING DATE: 1997-06-19
; PRIOR APPLICATION NUMBER: 60/022,999
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: 60/021,087
; PRIOR FILING DATE: 1996-07-02
; PRIOR APPLICATION NUMBER: 08/665,926
; PRIOR FILING DATE: 1996-06-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-215-224-3

Alignment Scores:
Pred. No.: 300 Length: 509
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 14 Gaps: 0
```

```
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 14 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-215-224-3 (1-509)
QY 1083 CGTCCCGAGCGTGGCCGCGTGCA 1106
Db 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 31
US-10-225-060-10
; Sequence 10, Application US/10225060
; Publication No. US20030092891A1
; GENERAL INFORMATION:
; APPLICANT: Davis et al.
; TITLE OF INVENTION: Expressed Ligand - Vascular Intercellular Signaling
; FILE REFERENCE: REG 333-Z
; CURRENT APPLICATION NUMBER: US/10/225,060
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US/09/709,188
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 08/740,223
; PRIOR FILING DATE: 1996-10-25
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-060-10

Alignment Scores:
Pred. No.: 300 Length: 509
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 14 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-225-060-10 (1-509)
QY 1083 CGTCCCGAGCGTGGCCGCGTGCA 1106
Db 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 32
US-09-833-745-57
; Sequence 57, Application US/09833745
; Patent No. US20020052038A1
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, JOSEPH
; APPLICANT: SETHURAMAN, NATHARAJAN
; APPLICANT: MACALLISTER, THOMAS
; TITLE OF INVENTION: CLONING, OVEREXPRESSION AND THERAPEUTIC USE OF
; FILE REFERENCE: BIOACTIVE HISTIDINE AMMONIA LYASE
; CURRENT APPLICATION NUMBER: US/09/833,745
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,770
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-833-745-57

Alignment Scores:
Pred. No.: 300 Length: 509
```

Pred. No.: 300 Length: 513
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 9 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-833-745-57 (1-513)

QY 137 AGCCATTATATCACTGCTCCA 160
DB 459 LysPProLeuLeuSerThrAlaPro 466

RESULT 33
US-10-424-599-231545
; Sequence 231545, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 231545
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51106C.1.pcp
US-10-424-599-231545

Alignment Scores:
Pred. No.: 299 Length: 529
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 12 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-424-599-231545 (1-529)

QY 374 GTGGTTCCTGAGTCTCTCGTT 397
DB 512 ValGlyLeuLeuSerSerLeuVal 519

RESULT 34
US-08-485-163-5
; Sequence 5, Application US/08485163
; Publication No. US20020098191A1
; GENERAL INFORMATION:
; APPLICANT: Beaudry, Gary A.
; APPLICANT: Maddon, Paul J.
; TITLE OF INVENTION: CD4-GAMMA2 CD4-IGG2 CHIMERAS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,163

FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-II-1-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: homo sapien
; CELL TYPE: lymphocyte
US-08-485-163-5

Alignment Scores:
Pred. No.: 299 Length: 530
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 0 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-485-163-5 (1-530)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
DB 78 AspArgAlaAspSerArgSer 85

RESULT 35
US-09-766-995-4
; Sequence 4, Application US/09766995
; Patent No. US20020052481A1
; GENERAL INFORMATION:
; APPLICANT: Graham P. Allaway et al.
; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED CD4-GAMMA2 AND CD4-IGG2 IMMUNOCON
; FILE REFERENCE: 2048/41215-CB/JPM/SHS
; CURRENT APPLICATION NUMBER: US/09/766,995
; CURRENT FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 530
; TYPE: PRT
; ORGANISM: homo sapians
US-09-766-995-4

Alignment Scores:
Pred. No.: 299 Length: 530
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 9 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-766-995-4 (1-530)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
DB 78 AspArgAlaAspSerArgSer 85

RESULT 36
US-09-939-537-6
; Sequence 6, Application US/09939537
; Publication No. US20030138410A1

GENERAL INFORMATION:
APPLICANT: Seed, Brian
Banapour, Babak
Romeo, Charles
Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/939,537
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-939-537-6
Alignment Scores:
Pred. No.: 299 Length: 532
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 10 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-939-537-6 (1-532)
QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
DB 78 AspArgAlaAspSerArgArgSer 85
RESULT 37
US-09-243-008-6
Sequence 6, Application US/09243008
Publication No. US20040005334A1
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Redirection of Cellular Immunity by
Receptor Chimeras

NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/243,008
FILING DATE: 02-Feb-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/394,176
FILING DATE: SEPTEMBER 11, 1995
APPLICATION NUMBER: 08/203,866
FILING DATE: February 28, 1994
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Karen F. Lech, Ph.D.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/270001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-243-008-6
Alignment Scores:
Pred. No.: 299 Length: 532
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 11 Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-243-008-6 (1-532)
QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
DB 78 AspArgAlaAspSerArgArgSer 85
RESULT 38
US-10-046-232-22
Sequence 22, Application US/10046232
Publication No. US20030119099A1
GENERAL INFORMATION:
APPLICANT: Helmut SCHWAB
APPLICANT: Anton GLEIDER
APPLICANT: Christoph KRATKY
APPLICANT: Ingrid DREVENY
APPLICANT: Peter POCHLAUER
APPLICANT: Wolfgang SKRANC
APPLICANT: Herbert MAYRHOFER
APPLICANT: Irma WIRTH
APPLICANT: Rudolf NEUHOFER
APPLICANT: Rodolfo BONA
TITLE OF INVENTION: New genes containing a DNA sequence coding for a hydroxynitrile 1
recombinant proteins derived therefrom and having hydroxynitrile

;; TITLE OF INVENTION: thereof
;; FILE REFERENCE: 2001-1882A/LC/01553
;; CURRENT APPLICATION NUMBER: US/10/046,232
;; CURRENT FILING DATE: 2002-10-31
;; PRIOR APPLICATION NUMBER: A60/2001
;; PRIOR FILING DATE: 2001-01-16
;; PRIOR APPLICATION NUMBER: A523/2001
;; PRIOR FILING DATE: 2001-04-03
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 22
;; LENGTH: 534
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of the artificial sequence: Hybrid protein PamHNL5xGC
US-10-046-232-22

Alignment Scores:
Pred. No.: 299 Length: 534
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 14 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-046-232-22 (1-534)

QY 386 AGTTCCTGCTGTTCCCTGACC 409
|||||
Db 7 SerSetLeuValValSerLeuAla 14

RESULT 39
US-10-156-761-11387
; Sequence 11387, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMTURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11387
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11387

Alignment Scores:
Pred. No.: 298 Length: 544
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 14 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-156-761-11387 (1-544)

QY 1648 ACCTTACCCCTTCCTGACACG 1671
|||||
Db 13 ThrSerThrProSerLeuThrSer 20

RESULT 40
US-10-628-088-415

;; Sequence 415, Application US/10628088
;; Publication No. US20040096451A1
;; GENERAL INFORMATION:
;; APPLICANT: Young, James
;; APPLICANT: Kiener, Peter
;; APPLICANT: Osterhaus, Albertus
;; APPLICANT: Fouchier, Ronaldus
;; TITLE OF INVENTION: METHODS OF TREATING AND PREVENTING
;; TITLE OF INVENTION: RSV, HMPV, AND PIV USING ANTI-RSV,
;; TITLE OF INVENTION: ANTI-HMPV, AND ANTI-PIV ANTIBODIES
;; FILE REFERENCE: 10271-072-999
;; CURRENT APPLICATION NUMBER: US/10/628,088
;; CURRENT FILING DATE: 2003-07-25
;; PRIOR APPLICATION NUMBER: 60/398,475
;; PRIOR FILING DATE: 2002-07-25
;; NUMBER OF SEQ ID NOS: 437
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 415
;; LENGTH: 555
;; TYPE: PRT
;; ORGANISM: Human parainfluenza virus 1 strain Washington/1964
;; FEATURE:
;; OTHER INFORMATION: F glycoprotein of Human parainfluenza 1 virus
US-10-628-088-415

Alignment Scores:
Pred. No.: 298 Length: 555
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 16 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-628-088-415 (1-555)

QY 1193 AGCTATTGCTGCTTCCTCCAGCCTG 1170
|||||
Db 13 SerLeuLeuSerSerSerSerLeu 20

RESULT 41
US-09-939-537-4
; Sequence 4, Application US/09939537
; Publication No. US20030138410A1
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; Banapur, Babak
; Romeo, Charles
; Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elding LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,537
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395

FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elding, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-939-537-4

Alignment Scores:
Pred. No.: 296 Length: 575
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 10 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-939-537-4 (1-575)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
Db 78 AsparGalaAspSerArgSer 85

RESULT 42

US-09-243-008-4
Sequence 4, Application US/09243008
Publication No. US20040005334A1
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Redirection of Cellular Immunity by
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/243,008
FILING DATE: 02-Feb-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/394,176
FILING DATE: SEPTEMBER 11, 1995
APPLICATION NUMBER: 08/203,866
FILING DATE: February 28, 1994
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Karen F. Lech, Ph.D
REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/270001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-243-008-4

Alignment Scores:
Pred. No.: 296 Length: 575
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 11 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-243-008-4 (1-575)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
Db 78 AsparGalaAspSerArgSer 85

RESULT 43

US-09-934-060A-13
Sequence 13, Application US/09934060A
Patent No. US20020155121A1
GENERAL INFORMATION:
APPLICANT: Devico, Anthony L.
APPLICANT: Fouts, Timothy R.
APPLICANT: Tuskan, Robert G.
TITLE OF INVENTION: VIRUS COAT PROTEIN/RECEPTOR CHIMERAS AND METHODS OF USE
FILE REFERENCE: 4115-144 CIP
CURRENT APPLICATION NUMBER: US/09/934,060A
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 09/684,026
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/158,321
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 590
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthesized construct
NAME/KEY: MISC FEATURE
LOCATION: (586)..(586)
OTHER INFORMATION: Xaa can be any amino acid
NAME/KEY: MISC FEATURE
LOCATION: (589)..(589)
OTHER INFORMATION: Xaa can be any amino acid
US-09-934-060A-13

Alignment Scores:
Pred. No.: 296 Length: 590
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 9 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-934-060A-13 (1-590)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
Db 450 AsparGalaAspSerArgSer 457

```
RESULT 44
US-10-062-937B-5
; Sequence 5, Application US/10062937B
; Publication No. US20030022195A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A. J.
; TITLE OF INVENTION: 59914 and 59921, CHOLINE TRANSPORTERS
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MP12001-005PIRM
; CURRENT APPLICATION NUMBER: US/10/062,937B
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/267,076
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FaSTSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-062-937B-5

Alignment Scores:
Pred. No.: 292          Length: 653
Score: 8.00           Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.24%          Indels: 0
DB: 14                 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-062-937B-5 (1-653)
QY 1493 TGGCGCGTGCCTGCTGCTGTT 1470
DB 564 TTPAlaValProLeuLeuVal 571

RESULT 45
US-10-391-399-73
; Sequence 73, Application US/10391399
; Publication No. US20030219806A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Curtis, Roy A. J.
; APPLICANT: Lora, Jose M.
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: NOVEL 18607, 15603, 69318, 12303, 48000,
; TITLE OF INVENTION: 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921
; TITLE OF INVENTION: AND 33751 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MP103-020MMIM
; CURRENT APPLICATION NUMBER: US/10/391,399
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/789,481
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: US 09/634,669
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: US 09/583,373
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: US 09/510,706
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 10/309,804
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: US 60/336,936
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 10/094,214
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US60/275,078
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/828,035
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
```

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FaSTSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-391-399-73

Alignment Scores:
Pred. No.: 292          Length: 653
Score: 8.00           Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.24%          Indels: 0
DB: 15                 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-391-399-73 (1-653)
QY 1493 TGGCGCGTGCCTGCTGCTGTT 1470
DB 564 TTPAlaValProLeuLeuVal 571

RESULT 46
US-09-833-745-39
; Sequence 39, Application US/09833745
; Patent No. US20020052038A1
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, JOSEPH
; APPLICANT: SETHURAMAN, NATARAJAN
; APPLICANT: MACALLISTER, THOMAS
; TITLE OF INVENTION: CLONING, OVEREXPRESSION AND THERAPEUTIC USE OF
; FILE REFERENCE: 078728/0106
; CURRENT APPLICATION NUMBER: US/09/833,745
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,770
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-833-745-39

Alignment Scores:
Pred. No.: 291          Length: 677
Score: 8.00           Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.20%          Indels: 0
DB: 9                 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-833-745-39 (1-677)
QY 137 AAGCATTATATTCACCTGCTCCA 160
DB 584 LysProLeuIleSerThrAlaPro 591

RESULT 47
US-10-369-493-6901
; Sequence 6901, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
```

```
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6901
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6901

Alignment Scores:
Pred. No.: 291          Length: 677
Score: 8.00            Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.20%          Indels: 0
DB: 15                  Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-369-493-6901 (1-677)

QY 137 AAGCATTATATCACTGCTCCA 160
Db 584 LysProLeuIleSerIlnAlaPro 591

RESULT 48
US-09-792-630-33
; Sequence 33, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/792,630
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 711
; TYPE: PRT
; ORGANISM: chipmunk parvovirus
US-09-792-630-33

Alignment Scores:
Pred. No.: 290          Length: 711
Score: 8.00            Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.24%          Indels: 0
DB: 9                  Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-792-630-33 (1-711)

QY 1647 CTGGGGGGCCAGATATCCGAGTG 1624
Db 404 LeuGIyGIyGlnAspIleArgVal 411

RESULT 49
US-09-953-351-33
; Sequence 33, Application US/09953351
; Publication No. US20030036643A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION I
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/953,351
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
```

```
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 711
; TYPE: PRT
; ORGANISM: chipmunk parvovirus
US-09-953-351-33

Alignment Scores:
Pred. No.: 290          Length: 711
Score: 8.00            Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.24%          Indels: 0
DB: 10                  Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-953-351-33 (1-711)

QY 1647 CTGGGGGGCCAGATATCCGAGTG 1624
Db 404 LeuGIyGIyGlnAspIleArgVal 411

RESULT 50
US-10-080-376-33
; Sequence 33, Application US/10080376
; Publication No. US20020172968A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Dahiyat, Basail I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295-2/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/080,376
; CURRENT FILING DATE: 2000-02-19
; PRIOR APPLICATION NUMBER: US 09/792,630
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 711
; TYPE: PRT
; ORGANISM: chipmunk parvovirus
US-10-080-376-33

Alignment Scores:
Pred. No.: 290          Length: 711
Score: 8.00            Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.24%          Indels: 0
DB: 13                  Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-10-080-376-33 (1-711)

QY 1647 CTGGGGGGCCAGATATCCGAGTG 1624
Db 404 LeuGIyGIyGlnAspIleArgVal 411

Search completed: June 21, 2004, 13:19:42
Job time : 166.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 13:05:47 ; Search time 24.5 Seconds
(without alignments)
8432.944 Million cell updates/sec

Title: US-10-029-345A-108_COPY_532_2532

Perfect score: 667

Sequence: 1 gctgcacatgccccatgagat.....tggaaatcatctgagtcctcc 2001

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Word size: 1

Total number of hits satisfying chosen parameters: 56994

Minimum DB seq length: 473

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlp
-Q/cg2_1/USPTO.spool.p/US10029345/funat.21062004.124616.5729/app.query.fasta_1.2183
-DB=Issued Patents AA -QFMT=fasta -SUFFIX=oligo.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=75 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEADSIZE=500 -MINLEN=473 -MAXLEN=200000000
-USER=US10029345 @CGN 1 1 27 @funat.21062004.124616.5729 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -OSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6
-FAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cg2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cg2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cg2_6/prodata/2/1aa/5A_COMB.pep:*
4: /cg2_6/prodata/2/1aa/5B_COMB.pep:*
5: /cg2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cg2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	70.8	665	4 US-09-816-494-2	Sequence 2, Appl1
2	10	1.6	503	4 US-09-134-000C-5978	Sequence 5978, Ap
3	3	1.4	737	4 US-08-989-299-6	Sequence 6, Appl1
4	4	1.4	737	4 US-09-407-427-6	Sequence 11, Appl1
5	5	1.2	503	3 US-08-740-223A-11	Sequence 11, Appl1
6	6	1.2	503	4 US-09-709-188-11	Sequence 11, Appl1
7	7	1.2	509	2 US-08-665-926-8	Sequence 8, Appl1
8	8	1.2	509	3 US-08-740-223A-10	Sequence 10, Appl1
9	9	1.2	509	4 US-09-202-491-2	Sequence 2, Appl1
10	10	1.2	509	4 US-09-202-491-3	Sequence 3, Appl1
11	11	1.2	509	4 US-09-709-188-10	Sequence 10, Appl1
12	12	1.2	511	4 US-09-252-991A-18691	Sequence 18691, A

13	8	1.2	530	3 US-08-477-460B-4	Sequence 4, Appl1
14	8	1.2	530	3 US-08-379-516-4	Sequence 4, Appl1
15	8	1.2	530	3 US-09-329-916-4	Sequence 4, Appl1
16	8	1.2	530	3 US-08-483-372A-4	Sequence 4, Appl1
17	8	1.2	530	4 US-09-409-006A-4	Sequence 4, Appl1
18	8	1.2	530	4 US-08-484-681-4	Sequence 4, Appl1
19	8	1.2	530	5 PCT-US93-07422-4	Sequence 4, Appl1
20	8	1.2	532	2 US-08-417-495-6	Sequence 6, Appl1
21	8	1.2	532	2 US-08-284-391B-6	Sequence 6, Appl1
22	8	1.2	532	2 US-09-218-950-6	Sequence 6, Appl1
23	8	1.2	532	5 PCT-US92-01785-6	Sequence 6, Appl1
24	8	1.2	532	5 PCT-US95-00454-6	Sequence 6, Appl1
25	8	1.2	575	2 US-08-417-495-4	Sequence 4, Appl1
26	8	1.2	575	2 US-08-284-391B-4	Sequence 4, Appl1
27	8	1.2	575	2 US-09-218-950-4	Sequence 4, Appl1
28	8	1.2	575	5 PCT-US92-01785-4	Sequence 4, Appl1
29	8	1.2	592	5 PCT-US95-00454-4	Sequence 4, Appl1
30	8	1.2	604	2 US-08-746-283-32	Sequence 18124, A
31	8	1.2	604	2 US-08-746-257A-30	Sequence 32, Appl1
32	8	1.2	604	2 US-08-746-257A-30	Sequence 30, Appl1
33	8	1.2	605	1 US-08-333-802-2	Sequence 2, Appl1
34	8	1.2	605	4 US-09-347-878-46	Sequence 46, Appl1
35	8	1.2	607	4 US-09-252-991A-26137	Sequence 26137, A
36	8	1.2	611	4 US-09-252-991A-28230	Sequence 28230, A
37	8	1.2	612	2 US-08-746-283-31	Sequence 31, Appl1
38	8	1.2	612	2 US-08-746-257A-29	Sequence 29, Appl1
39	8	1.2	613	2 US-08-746-283-1	Sequence 1, Appl1
40	8	1.2	613	2 US-08-746-283-3	Sequence 1, Appl1
41	8	1.2	613	2 US-08-746-283-5	Sequence 5, Appl1
42	8	1.2	613	2 US-08-746-257A-1	Sequence 1, Appl1
43	8	1.2	613	4 US-09-347-878-48	Sequence 48, Appl1
44	8	1.2	630	4 US-08-472-888A-6	Sequence 6, Appl1
45	8	1.2	683	4 US-09-591-095-6	Sequence 6, Appl1
46	8	1.2	903	1 US-08-021-601-12	Sequence 12, Appl1
47	8	1.2	903	1 US-08-082-849B-12	Sequence 12, Appl1
48	8	1.2	903	5 PCT-US94-01624-12	Sequence 12, Appl1
49	8	1.2	935	4 US-09-252-991A-33037	Sequence 33037, A
50	8	1.2	953	4 US-09-252-991A-27230	Sequence 27230, A
51	8	1.2	2500	2 US-08-801-263A-2	Sequence 2, Appl1
52	8	1.2	2500	3 US-09-102-248-2	Sequence 2, Appl1
53	8	1.2	2500	4 US-09-367-764-2	Sequence 9, Appl1
54	8	1.2	2512	2 US-08-801-263A-9	Sequence 9, Appl1
55	8	1.2	2512	3 US-09-102-248-9	Sequence 9, Appl1
56	8	1.2	2512	4 US-09-367-764-9	Sequence 9, Appl1
57	8	1.2	2517	2 US-08-801-263A-5	Sequence 5, Appl1
58	8	1.2	2517	3 US-09-102-248-5	Sequence 5, Appl1
59	8	1.2	2517	4 US-09-367-764-5	Sequence 5, Appl1
60	8	1.2	2842	1 US-07-741-940-7	Sequence 7, Appl1
61	8	1.2	2842	1 US-08-289-548A-7	Sequence 7, Appl1
62	8	1.2	2842	1 US-08-452-654-7	Sequence 7, Appl1
63	8	1.2	2842	1 US-08-452-654-7	Sequence 7, Appl1
64	8	1.2	2843	1 US-07-741-940-2	Sequence 2, Appl1
65	8	1.2	2843	1 US-08-289-548A-2	Sequence 2, Appl1
66	8	1.2	2843	1 US-08-452-654-2	Sequence 2, Appl1
67	8	1.2	2843	1 US-08-452-655B-2	Sequence 2, Appl1
68	8	1.2	2843	1 US-08-452-655B-2	Sequence 2, Appl1
69	8	1.2	2843	2 US-08-370-235A-2	Sequence 2, Appl1
70	8	1.2	2843	3 US-08-450-582-2	Sequence 2, Appl1
71	8	1.2	2843	3 US-08-450-582-7	Sequence 2, Appl1
72	8	1.2	2843	4 US-08-449-731-2	Sequence 2, Appl1
73	8	1.2	2973	2 US-08-821-355A-7	Sequence 7, Appl1
74	8	1.2	2973	2 US-09-003-687A-7	Sequence 7, Appl1
75	8	1.2	2973	3 US-09-136-605-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.

;; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
;; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
;; FILE REFERENCE: 10448-030002
;; CURRENT APPLICATION NUMBER: US/09/816,494
;; PRIORITY FILING DATE: 2001-03-23
;; PRIOR APPLICATION NUMBER: US 60/191,858
;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 665
;; TYPE: prt
;; ORGANISM: Homo sapiens
US-09-816-494-2

Alignment Scores:

Pred. No.:	0	Length:	665
Score:	472.00	Matches:	472
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	70.76%	Indels:	0
	4	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-816-494-2 (1-665)

QY 586 ACCTGTCCTCAAGCTGATCTTATCCCGAGTCTCATTTCTGCGTGTGATGAC 645
Db 194 ThrcysprolyspromaphellieProgluserhisheleuargvalProvalaasp 213
QY 646 AGCTTTGTGAGAAATTTTCCCGTGTGACCAATCAGTATGATTTCATTTGAGAAACA 705
Db 214 SerphecylulysileuProtripleuaspyservalaspheileglulysala 233
QY 706 AAAGCTTCAATGATGTGTCTAGTGCATGTTTGTGAGTCCCGCTCCGACAC 765
Db 234 LyelaserhenglycysvalleuvalhisCysleuvalagilylleserargseralatr 253
QY 766 ATGCTTATGCTTCATCATGAAAGAGATGACATGCTTTAGATGAAGCTTACAGATT 825
Db 254 lilelalelalyrilemetlysargmetaspmetserleuaspglulalatyargphe 273
QY 826 GTGAAAGAAAAGACCTACTATATCTCCAACTTCAATTTTGTGGCCCACTCTGGAC 885
Db 274 VallysglulysargprohmrlieserProasnpheasnpheleuylglinleuasp 293
QY 886 TATGAGAAGAAATTAAGAACAGACTGAGCATCAGGCGCAAAAGCAACTCAAGCTG 945
Db 294 TyrglulyselylleylsanglinthrglylaserdlyProlyserlyseuylseu 313
QY 946 CTGCACCTGAGAAAGCAATGAACCTGTCTCTGCTCTCAAGGGTGGACAGAAAGC 1005
Db 314 LeuhisleuglulysProasngluprovalProalavalsergluglyglinlysser 333
QY 1006 GAGACGCCCTCAGTGCACCTGTGCGCACTGCGCACTCGAGGAGGAGAGAAAG 1065
Db 334 GluthrProleuserProcyosalaaapseralathmserglulalalalyindrg 353
QY 1066 CCCGTCATCCCGCAGCGCTGCCAGCGTCCAGCGTGCACCGCTCGCTTTAGAGAC 1125
Db 354 ProvalhsproalaserValProserValGlnProserleuenglunsp 373
QY 1126 ACCCGCTGTACAGCGCGCTGAGTGGCGGCACTGTCCGAGACAGCGCTGAGACAGC 1185
Db 374 SerProleuValGlnlaleuserglyleuhsleuseralaaaspargleuaspser 393
QY 1186 AATAAGCTCAAGCTTCTTCTCTGATATCAATCAGTTTCATTTAGCCAGCATG 1245
Db 394 AenlyseuLyArgserPheuserleuaspilleysservalserlyseralaseret 413
QY 1246 GAGAGATCTTACATGAGCTTCTCTCATCAAGAAAGTGTGAAATATCAAACTTCC 1305
Db 414 AlaalaaserleuhsrglyPheuserserserglunspalalenglulutyryllyseProser 433

QY 1306 ACTACTCTGATGGAGCAACAAGCTATGCCAGTTCTCCCTGTTCCAGAACTATCGAG 1365
Db 434 ThrtThreleuaspolythreanlyseuysglinpneserProvalglululeuserglu 453
QY 1366 CAGACTCCCGAAACCAAGTCTCTGATAGAGGAGAAAGCAGATATCCCAAGAGCTGACAGC 1425
Db 454 GluthrProglulnhsersProaspilysglulualaserlleProlyseylseuInthr 473
QY 1426 GCCAGGCTTTCAGACAGCAGCAGCAAGCATGTCATTCGTGAGAACAGCAGCATGGC 1485
Db 474 AlaargProserAspsercinserlysargleuhsleuseralargthnserseuergly 493
QY 1486 ACCGCCCAAGGTCCTTTTATCTTCACATGCAATGAGTGGAGCGTGGAGACATTAAC 1545
Db 494 ThrtAlaglnArgserleuuserProleuhsArgsergllyseuValGluaspasnlyr 513
QY 1546 CACACAGCTTCTTTTGGGCTTTCCACACGCGCAGCAGCAGCTTACAGAACTGTGCTGC 1605
Db 514 HsthrserPheleuPhegllyseuuserThrserglnghnhsleuthrlysserlalgly 533
QY 1606 CTGGGCTTAAAGGCTGCGCACTCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTG 1665
Db 534 LeuglyleuLygLyTtrpHisseraspilleuvalaroglinthnserThnProserleu 553
QY 1666 ACCAGCAGCTGGTATTTTGGCCACAGAGTCTTCACTTCTACTGTGCTCAGCCATCTAC 1725
Db 554 Thrserserlyrlyrphalathrgluserserhispheryrseralaserlalelyr 573
QY 1726 GGAGCAGTGCAGATTACTGTGCTTACAGCTGACAGCAGCGCCACCTTGGCGAGCA 1785
Db 574 GlylyseralaserlyrseralalyrserCysserglneupromtrCysglyaspGln 593
QY 1786 GTCTTTTGTGCGCAGGCGCGCAGAACGCAAGTCAAGAGCTGACTGCGCGAGAGCTGG 1845
Db 594 ValtyrserValargargargglulysProseraspbrglaaapserargargserTtrp 613
QY 1846 CATGAAGAGACCCCTTTGAAAAGCAGTTTAAAGCGAAGCTGCCAAATGGAATTTGGA 1905
Db 614 HlsglulnserProphleglulysglnphelysartrgserCysglnmetglupnegly 633
QY 1906 GAGAGCATCATGTACAGAAACAGGTCAAGGAGAGAGCTGGGAAAGTGGGAGCTGACT 1965
Db 634 Gluserilemetsergluamargserarggluenglulysvalglyserglnser 653
QY 1966 AGCTTTCCGGCAGCATGAAATCATTTAGAGCTTCC 2001
Db 654 Serpheserglysermetglulleleglulalser 665

RESULT 2

US-09-134-000C-5978
; Sequence 5978, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucelte-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIORITY FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5978
; LENGTH: 503
; TYPE: prt
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5978

Alignment Scores:

Pred. No.:	0	Length:	503
Score:	515	Matches:	10
Percent Similarity:	100.00%	Conservative:	0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.55% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-134-000C-5978 (1-503)

QY 103 GCCGCTATCATATTAGCAGCAGCTTTTCCG 74
DB 347 AAGlYTYrGlnleuAlaAlaIeuphePro 356

RESULT 3
US-08-989-299-6
; Sequence 6, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Robinson, Keith E.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; City: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E. Beth
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-025.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-989-299-6

Alignment Scores:
Pred. No.: 4.6 Length: 737
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: 3 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-989-299-6 (1-737)

QY 1334 CATAGCTGTGGTCCCATCCAGAGTA 1308
DB 22 HisserleuValProserArgVal 30

RESULT 4
US-09-407-427-6
; Sequence 6, Application US/09407427
; Patent No. 6610497
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Robinson, Keith E.

; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG AND THERAPEUTIC
; TITLE OF INVENTION: AND DIAGNOSTIC USES THEREFOR
; FILE REFERENCE: MNT-132CP2
; CURRENT APPLICATION NUMBER: US/09/407,427
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 09/163,648
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 08/989,299
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; US-09-407-427-6

Alignment Scores:
Pred. No.: 4.6 Length: 737
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-407-427-6 (1-737)

QY 1334 CATAGCTGTGGTCCCATCCAGAGTA 1308
DB 22 HisserleuValProserArgVal 30

RESULT 5
US-08-740-223A-11
; Sequence 11, Application US/08740223A
; Patent No. 6265564
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: Expressed ligand - Vascular
; TITLE OF INVENTION: Inter cellular Signalling Molecule
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill Road
; City: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/740,223A
; FILING DATE: 25-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: USSN 60/022/999
; FILING DATE: 02-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Cobert, Robert J
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 333
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```
/
/ FEATURE:
/ NAME/KEY: mtl3
/ LOCATION: 1...503
/ OTHER INFORMATION: mouse TIE ligand-3
US-08-740-223A-11

Alignment Scores:
Pred. No.: 44.5 Length: 503
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 3 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-740-223A-11 (1-503)
QY 1083 CGTGGCCAGCGTGGCCAGCGTGCA 1106
DB 80 ArgAlaGlnArgAlaGlnArgAla 87

RESULT 6
US-09-709-188-11
/ Sequence 11, Application US/09709188
/ Patent No. 6441137
/ GENERAL INFORMATION:
/ APPLICANT: Davis et al.
/ TITLE OF INVENTION: Expressed Ligand - Vascular Intercellular Signaling Molecule
/ FILE REFERENCE: REG 333-2
/ CURRENT APPLICATION NUMBER: US/09/709,188
/ PRIOR FILING DATE: 2000-11-09
/ PRIOR APPLICATION NUMBER: 08/740,223
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 11
/ LENGTH: 503
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-709-188-11

Alignment Scores:
Pred. No.: 44.5 Length: 503
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-709-188-11 (1-503)
QY 1083 CGTGGCCAGCGTGGCCAGCGTGCA 1106
DB 80 ArgAlaGlnArgAlaGlnArgAla 87

RESULT 7
US-08-665-926-8
/ Sequence 8, Application US/08665926
/ Patent No. 5851797
/ GENERAL INFORMATION:
/ APPLICANT: Valenzuela et al.
/ TITLE OF INVENTION: TIE LIGAND-3, METHODS OF MAKING AND USES
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Regeneron Pharmaceuticals, Inc.
/ STREET: 777 Old Saw Mill River Road
/ CITY: Tarrytown
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10591-6707
/ TELEPHONE: 914-345-7400
/ TELEFAX: 914-345-7721
/ INFORMATION FOR SEQ ID NO: 10:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
```

```
/
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/665,926
/ FILING DATE: 19-JUN-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Robert J. Covert
/ REGISTRATION NUMBER: 36,108
/ REFERENCE/DOCKET NUMBER: REG 330-H
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (914) 345-7400
/ TELEFAX: (914) 345-2113
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 509 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-665-926-8

Alignment Scores:
Pred. No.: 44.4 Length: 509
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-665-926-8 (1-509)
QY 1083 CGTGGCCAGCGTGGCCAGCGTGCA 1106
DB 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 8
US-08-740-223A-10
/ Sequence 10, Application US/08740223A
/ Patent No. 6265564
/ GENERAL INFORMATION:
/ APPLICANT: Davis, et al.
/ TITLE OF INVENTION: Expressed Ligand - Vascular
/ TITLE OF INVENTION: Intercellular Signalling Molecule
/ NUMBER OF SEQUENCES: 28
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Regeneron Pharmaceuticals, Inc.
/ STREET: 777 Old Saw Mill Road
/ CITY: Tarrytown
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10591
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/740,223A
/ FILING DATE: 25-OCT-1996
/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: USSN 60/022/999
/ FILING DATE: 02-AUG-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Covert, Robert J
/ REGISTRATION NUMBER: 36,108
/ REFERENCE/DOCKET NUMBER: REG 333
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 914-345-7400
/ TELEFAX: 914-345-7721
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 509 amino acids
```



```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: TIE ligand-3
; LOCATION: 1...509
; OTHER INFORMATION:
US-08-740-223a-10

Alignment Scores:
Pred. No.: 44.4
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.20%
DB: 3
Length: 509
Matches: 8
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-08-740-223a-10 (1-509)
QY 1083 CGTCCCGAGCGTCCCGCGGTGCA 1106
DB 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 9
US-09-202-491-2
; Sequence 2, Application US/09202491
; Patent No. 643267
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
; FILE REFERENCE: REG310-K
; CURRENT APPLICATION NUMBER: US/09/202,491
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: PCT/US97/10728
; EARLIER FILING DATE: 1997-06-19
; EARLIER APPLICATION NUMBER: 60/022,999
; EARLIER FILING DATE: 1996-08-02
; EARLIER APPLICATION NUMBER: 60/021,087
; EARLIER FILING DATE: 1996-07-02
; EARLIER APPLICATION NUMBER: 08/665,926
; EARLIER FILING DATE: 1996-06-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-202-491-2

Alignment Scores:
Pred. No.: 44.4
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.20%
DB: 4
Length: 509
Matches: 8
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-09-202-491-2 (1-509)
QY 1083 CGTCCCGAGCGTCCCGCGGTGCA 1106
DB 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 10
US-09-202-491-3
; Sequence 3, Application US/09202491
; Patent No. 643267
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
; FILE REFERENCE: REG310-K
```

```
; CURRENT APPLICATION NUMBER: US/09/202,491
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: PCT/US97/10728
; EARLIER FILING DATE: 1997-06-19
; EARLIER APPLICATION NUMBER: 60/022,999
; EARLIER FILING DATE: 1996-08-02
; EARLIER APPLICATION NUMBER: 60/021,087
; EARLIER FILING DATE: 1996-07-02
; EARLIER APPLICATION NUMBER: 08/665,926
; EARLIER FILING DATE: 1996-06-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-202-491-3

Alignment Scores:
Pred. No.: 44.4
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.20%
DB: 4
Length: 509
Matches: 8
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-09-202-491-3 (1-509)
QY 1083 CGTCCCGAGCGTCCCGCGGTGCA 1106
DB 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 11
US-09-709-188-10
; Sequence 10, Application US/09709188
; Patent No. 644137
; GENERAL INFORMATION:
; APPLICANT: Davis et al.
; TITLE OF INVENTION: Expressed Ligand - Vascular Intercellular Signaling Molecule
; FILE REFERENCE: REG 333-Z
; CURRENT APPLICATION NUMBER: US/09/709,188
; EARLIER FILING DATE: 2000-11-09
; EARLIER APPLICATION NUMBER: 08/740,223
; PRIOR FILING DATE: 1996-10-25
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-709-188-10

Alignment Scores:
Pred. No.: 44.4
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.20%
DB: 4
Length: 509
Matches: 8
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-09-709-188-10 (1-509)
QY 1083 CGTCCCGAGCGTCCCGCGGTGCA 1106
DB 86 ArgAlaGlnArgAlaGlnArgAla 93

RESULT 12
US-09-252-991a-18691
; Sequence 18691, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
```

;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; PRIOR FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 18691
;; LENGTH: 511
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18691

Alignment Scores:
Pred. No.: 44.4 Length: 511
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-252-991A-18691 (1-511)

QY 434 TCCCTACTGCATTTCTAGCCTT 457
Db |||||
406 SerLeuProAlaPheLeuSerLeu 413

RESULT 13
US-08-477-460B-4

;; Sequence 4, Application US/08477460B
;; Patent No. 6034223

;; GENERAL INFORMATION:
;; APPLICANT: Progenics Pharmaceuticals, Inc.
;; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
;; NUMBER OF SEQUENCES: 9
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Cooper & Dunham
;; STREET: 30 Rockefeller Plaza
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10112

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.24
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/477,460B
;; FILING DATE: 07-JUN-1995

;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/927,931
;; FILING DATE: 07-AUG-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: White, John P.
;; REGISTRATION NUMBER: 28,678
;; REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/AJM
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 977-9550
;; TELEFAX: (212) 977-9809
;; TELEX: 422523 COOP UI
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 530 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: CDNA
;; ORIGINAL SOURCE:

;; ORGANISM: homo sapien
;; CELL TYPE: Lymphocyte
US-08-477-460B-4

Alignment Scores:
Pred. No.: 44.2 Length: 530
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 3 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-477-460B-4 (1-530)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
Db |||||
78 AspArgAlaAspSerArgArgSer 85

RESULT 14

US-08-379-516-4
;; Sequence 4, Application US/08379516
;; Patent No. 6083478

;; GENERAL INFORMATION:
;; APPLICANT: Allway, Graham P.
;; TITLE OF INVENTION: No. 6083478-Peptidyl Moiety-Conjugated CD4-Gamma2 and CD4-IgG2
;; FILE REFERENCE: 41215-A-PCT-US
;; CURRENT APPLICATION NUMBER: US/08/379,516
;; CURRENT FILING DATE: 1996-06-10
;; EARLIER APPLICATION NUMBER: PCT/US93/07422
;; EARLIER FILING DATE: 1993-08-06
;; EARLIER APPLICATION NUMBER: 07/927,931
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 530
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-08-379-516-4

Alignment Scores:
Pred. No.: 44.2 Length: 530
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 3 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-379-516-4 (1-530)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
Db |||||
78 AspArgAlaAspSerArgArgSer 85

RESULT 15
US-09-329-916-4

;; Sequence 4, Application US/09329916
;; Patent No. 6177549
;; GENERAL INFORMATION:
;; APPLICANT: Progenics Pharmaceuticals, Inc.
;; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
;; NUMBER OF SEQUENCES: 9
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Cooper & Dunham
;; STREET: 30 Rockefeller Plaza
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10112

;; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/329,916
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,460
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-329-916-4

Alignment Scores:
Pred. No.: 44.2 Length: 530
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 3 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-329-916-4 (1-530)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842

DB 78 AspArgAlaAspSerArgArgSer 85

RESULT 16
US-08-485-372A-4
Sequence 4, Application US/08485372A
Patent No. 6187748
GENERAL INFORMATION:
APPLICANT: Beaudry, Gary A.
APPLICANT: Maddon, Paul J.
TITLE OF INVENTION: CD4-GAMMA2 CD4-IgG2 CHIMERAS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,372A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,227
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-II-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-485-372A-4

Alignment Scores:
Pred. No.: 44.2 Length: 530
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 3 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-485-372A-4 (1-530)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842

DB 78 AspArgAlaAspSerArgArgSer 85

RESULT 17
US-09-409-006A-4
Sequence 4, Application US/09409006A
Patent No. 6342586
GENERAL INFORMATION:
APPLICANT: Progenice Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/409,006A
FILING DATE: 29-SEP-1999
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: unknown
ORIGINAL SOURCE: cDNA
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-409-006A-4

Alignment Scores:

Pred. No.:	44.2	Length:	530
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-409-006A-4 (1-530)

QY 1819 GACAGAGCTGACTCGCGCGAGC 1842

Db 78 AsparGAlaAspSerArgArgSer 85

RESULT 18

US-08-484-681-4
Sequence 4, Application US/08484681
Patent No. 6451313
GENERAL INFORMATION:
APPLICANT: Beaudry, Gary A.
APPLICANT: Madden, Paul J.
TITLE OF INVENTION: CD4-GAMMA2 CD4-IgG2 CHIMERAS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,681
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-11-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-484-681-4

Alignment Scores:

Pred. No.:	44.2	Length:	530
Score:	8.00	Matches:	8

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.20%
DB: 4
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-484-681-4 (1-530)

QY 1819 GACAGAGCTGACTCGCGCGAGC 1842

Db 78 AsparGAlaAspSerArgArgSer 85

RESULT 19

PCT-US93-07422-4
Sequence 4, Application PC/TUS9307422
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOLETY-CONJUGATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07422
FILING DATE: 19930806
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/JJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UT
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
PCT-US93-07422-4

Alignment Scores:

Pred. No.:	44.2	Length:	530
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x PCT-US93-07422-4 (1-530)

QY 1819 GACAGAGCTGACTCGCGCGAGC 1842

Db 78 AsparGAlaAspSerArgArgSer 85

RESULT 20

US-08-417-495-6
; Sequence 6, Application US/08417495
; Patent No. 5843728
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Redirection of Cellular Immunity by
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,495
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,866
; FILING DATE:
; APPLICATION NUMBER: US/07/847,566
; FILING DATE:
; APPLICATION NUMBER: 07/665,961
; FILING DATE: March 7, 1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/119002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-417-495-6

Alignment Scores:
Pred. No.: 44.2 Length: 532
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 2 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-08-417-495-6 (1-532)
QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
Db 78 AspaRgAlaAspaSerArGaRgSer 85

RESULT 21
US-08-284-391B-6
; Sequence 6, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS

NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391B
; FILING DATE: 02-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-284-391B-6

Alignment Scores:
Pred. No.: 44.2 Length: 532
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 2 Gaps: 0

US-10-029-345a-108_COPY_532_2532 (1-2001) x US-08-284-391B-6 (1-532)
QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
Db 78 AspaRgAlaAspaSerArGaRgSer 85

RESULT 22
US-09-218-950-6
; Sequence 6, Application US/09218950
; Patent No. 6284240
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA

ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-6
Alignment Scores:
Pred. No.: 44.2
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.20%
Matches: 532
Conservative: 8
Mismatch: 0
Indels: 0
Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-218-950-6 (1-532)
QY 1819 GACAGAGCTGACTGCGCGGAGC 1842
DB 78 AsparGAlaAspSerArgArgSer 85
RESULT 23
PCT-US92-01785-6
Sequence 6, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-01785-6
Alignment Scores:
Pred. No.: 44.2
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.20%
Matches: 532
Conservative: 8
Mismatch: 0
Indels: 0
Gaps: 0
US-10-029-345A-108_COPY_532_2532 (1-2001) x PCT-US92-01785-6 (1-532)
QY 1819 GACAGAGCTGACTGCGCGGAGC 1842
DB 78 AsparGAlaAspSerArgArgSer 85
RESULT 24
PCT-US95-00454-6
Sequence 6, Application PC/TUS9500454
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00454-6

Alignment Scores:
Pred. No.: 44.2 Length: 532
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x PCT-US95-00454-6 (1-532)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
Db 78 AspArgAlaAspSerArgSer 85

RESULT 25
US-08-417-495-4
Sequence 4, Application US/08417495
Patent No. 5843728
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Redirection of Cellular Immunity by
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,495
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/203,866
FILING DATE:
APPLICATION NUMBER: US/07/847,566
FILING DATE:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-417-495-4

Alignment Scores:
Pred. No.: 43.9 Length: 575
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-417-495-4 (1-575)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842
Db 78 AspArgAlaAspSerArgSer 85

RESULT 26
US-08-284-391B-4
Sequence 4, Application US/08284391B
Patent No. 5851828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Ebling LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Ebling, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-4

Alignment Scores:
Pred. No.: 43.9 Length: 575
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-284-391B-4 (1-575)

QY 1819 GACAGAGCTGACTCGCGGAGC 1842
DB 78 AspaRgAlaSpSerArGArSer 85

RESULT 27

US-09-218-950-4
Sequence 4, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-4

Alignment Scores:

Pred. No.:	43.9	Length:	575
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	3	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-218-950-4 (1-575)

QY 1819 GACAGAGCTGACTCGCGGAGC 1842

|||||

DB 78 AspaRgAlaSpSerArGArSer 85

RESULT 28

PCT-US92-01785-4
Sequence 4, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
TITLE OF INVENTION: Chimeras
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: WordPerfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-01785-4

Alignment Scores:

Pred. No.:	43.9	Length:	575
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	5	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x PCT-US92-01785-4 (1-575)

QY 1819 GACAGAGCTGACTCGCGGAGC 1842

DB 78 AspaRgAlaSpSerArGArSer 85

RESULT 29

PCT-US95-00454-4
Sequence 4, Application PC/TUS9500454
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
TITLE OF INVENTION: Bearing Cells
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA

|||||

COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Mordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00454-4

Alignment Scores:
Pred. No.: 43.9 Length: 575
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x PCT-US95-00454-4 (1-575)

QY 1819 GACAGAGCTGACTGCGCGGAGC 1842
Db 78 AspArgLalaPseSerArgSer 85

RESULT 30
US-09-252-991A-18124
Sequence 18124, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18124
LENGTH: 592
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18124

Alignment Scores:
Pred. No.: 43.7 Length: 592

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-252-991A-18124 (1-592)

QY 1113 CGACGGCTGACGCTGCGGACGCT 1090
Db 302 ArgArgLeuAlaAlaGlyHisAla 309

RESULT 31
US-08-746-283-32
Sequence 32, Application US/08746283
Patent No. 5834280
GENERAL INFORMATION:
APPLICANT: Okenbolil, Karen M.
APPLICANT: St. Joan O.
TITLE OF INVENTION: Glucose Oxidases
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSER: No. 58342800 No. 5834280disk of No. 5834280th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,283
FILING DATE: 07-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Agria, Dr. Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 4158.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 867-0123
TELEFAX: (212) 878-9655
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 604 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-746-283-32

Alignment Scores:
Pred. No.: 43.6 Length: 604
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-746-283-32 (1-604)

QY 386 AGTTCCTGCTGTTGTTCCCTGCGC 409
Db 7 SerSerLeuValValSerLeuAla 14

RESULT 32
US-08-746-257A-30
Sequence 30, Application US/08746257A
Patent No. 5879921
GENERAL INFORMATION:
APPLICANT: Cherry, Joel

```

;
; APPLICANT: Berka, Randy
; APPLICANT: Halkier, Torben
; TITLE OF INVENTION: Glucose Oxidases
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 58799210 No. 5879921disk of No. 5879921ch America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/746,257A
; FILING DATE: 07-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Starnes, Robert L.
; REGISTRATION/DOCKET NUMBER: 41,324
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
;
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 604 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-746-257A-30
;
; Alignment Scores:
; Pred. No.: 43.6 Length: 604
; Score: 8.00 Matches: 8
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 1.20% Indels: 0
; DB: 2 Gaps: 0
;
; US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-746-257A-30 (1-604)
;
; QY 386 AGTTCCTCGTTGTTCCCTGGCC 409
;
; Db 7 SerSerIeuValValSerIeuAla 14
;
; RESULT 33
; US-08-333-802-2
; Sequence 2, Application US/08333802
; Patent No. 5516671
;
; GENERAL INFORMATION:
; APPLICANT: Lawrence, Ellen B.
; APPLICANT: Levine, Elaine B.
; APPLICANT: Shah, Dilipkumar M.
; TITLE OF INVENTION: Method of Controlling Plant Pathogens
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BBAF
; STREET: 700 Chesterfield Parkway No. 5516671ch
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

;
; APPLICATION NUMBER: US/08/333,802
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161041
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(10663)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6047
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 605 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
;
; US-08-333-802-2
;
; Alignment Scores:
; Pred. No.: 43.6 Length: 605
; Score: 8.00 Matches: 8
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 1.20% Indels: 0
; DB: 1 Gaps: 0
;
; US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-333-802-2 (1-605)
;
; QY 386 AGTTCCTCGTTGTTCCCTGGCC 409
;
; Db 7 SerSerIeuValValSerIeuAla 14
;
; RESULT 34
; US-09-347-878-46
; Sequence 46, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 46
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Aspergillus niger
;
; US-09-347-878-46
;
; Alignment Scores:
; Pred. No.: 43.6 Length: 605
; Score: 8.00 Matches: 8
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 1.20% Indels: 0
; DB: 4 Gaps: 0
;
; US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-347-878-46 (1-605)
;
; QY 386 AGTTCCTCGTTGTTCCCTGGCC 409
;
; Db 7 SerSerIeuValValSerIeuAla 14
;
; RESULT 35
; US-09-252-991A-26137
; Sequence 26137, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

```

```
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 26137
/ LENGTH: 607
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26137

Alignment Scores:
Pred. No.: 43.6 Length: 607
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
Gaps: 0
DB: 4

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-252-991A-26137 (1-607)
QY 1622 CAGCCCTTAAGCCCGAGCCAGCA 1599
DB 193 GlnProLeuArgProArgProAla 200

RESULT 36
US-09-252-991A-28230
/ Sequence 28230, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 28230
/ LENGTH: 611
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28230

Alignment Scores:
Pred. No.: 43.6 Length: 611
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
Gaps: 0
DB: 4

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-252-991A-28230 (1-611)
QY 1158 CCGTCCGAGACAGGCTGGAAGA 1181
DB 360 ProValArgArgGlnAlaGlyArg 367

RESULT 37
US-08-746-283-31
/ Sequence 31, Application US/08746283
/ Patent No. 5834280
/ GENERAL INFORMATION:
/ APPLICANT: Oxenboll, Karen M.
/ APPLICANT: Si, Joan Q.
```

```
/ APPLICANT: Aagaard, Jesper
/ TITLE OF INVENTION: Glucose Oxidases
/ NUMBER OF SEQUENCES: 32
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: No. 58342800 No. 5834280disk of No. 5834280th America
/ STREET: 405 Lexington Avenue
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10174
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Agtis, Dr. Cheryl H.
/ REGISTRATION NUMBER: 34,086
/ REFERENCE/DOCKET NUMBER: 4158,214-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 867-0123
/ TELEFAX: (212) 878-9655
/ INFORMATION FOR SEQ ID NO: 31:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 612 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: No. 5834280e
US-08-746-283-31

Alignment Scores:
Pred. No.: 43.6 Length: 612
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
Gaps: 0
DB: 2

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-746-283-31 (1-612)
QY 1593 CGTAGGTCTGCTGCTGTGGA 1570
DB 307 ArgGluValLeuLeuAlaGlyGly 314

RESULT 38
US-08-746-257A-29
/ Sequence 29, Application US/08746257A
/ Patent No. 5879921
/ GENERAL INFORMATION:
/ APPLICANT: Cherry, Joel
/ APPLICANT: Berka, Randy
/ APPLICANT: Halkier, Torben
/ TITLE OF INVENTION: Glucose Oxidases
/ NUMBER OF SEQUENCES: 30
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: No. 58799210 No. 5879921disk of No. 5879921ch America
/ STREET: 405 Lexington Avenue
/ CITY: New York
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10174
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/746,257A
```

FILING DATE: 07-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Starnes, Robert L.
REGISTRATION NUMBER: 41,334
REFERENCE/DOCKET NUMBER: 4700.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 612 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-746-257A-29

Alignment Scores:
Pred. No.: 43.6 Length: 612
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-746-257A-29 (1-612)

QY 1593 CGTGAGTGTCTGCTGCTGTGGA 1570
DB 307 ArggluValIleuLeuAlaIglyIy 314

RESULT 39
US-08-746-283-1
Sequence 1, Application US/08746283
Patent No. 5834280
GENERAL INFORMATION:
APPLICANT: Oxenboll, Karen M.
APPLICANT: Si, Joan O.
APPLICANT: Aagaard, Jesper
TITLE OF INVENTION: Glucose Oxidases
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 58342800 No. 5834280disk of No. 5834280th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,283
FILING DATE: 07-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Agtis, Dr. Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 4158.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 867-0123
TELEFAX: (212) 878-9655
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 613 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-746-283-1

Alignment Scores:
Pred. No.: 43.6 Length: 613
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-746-283-1 (1-613)

QY 1593 CGTGAGTGTCTGCTGCTGTGGA 1570
DB 308 ArggluValIleuLeuAlaIglyIy 315

RESULT 40
US-08-746-283-3
Sequence 3, Application US/08746283
Patent No. 5834280
GENERAL INFORMATION:
APPLICANT: Oxenboll, Karen M.
APPLICANT: Si, Joan O.
APPLICANT: Aagaard, Jesper
TITLE OF INVENTION: Glucose Oxidases
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 58342800 No. 5834280disk of No. 5834280th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,283
FILING DATE: 07-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Agtis, Dr. Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 4158.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 867-0123
TELEFAX: (212) 878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 613 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-746-283-3

Alignment Scores:
Pred. No.: 43.6 Length: 613
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-746-283-3 (1-613)

QY 1593 CGTGAGTGTCTGCTGCTGTGGA 1570
DB 308 ArggluValIleuLeuAlaIglyIy 315

RESULT 41
US-08-746-283-5

Sequence 5, Application US/08746283
Patent No. 5834280
GENERAL INFORMATION:
APPLICANT: Oxendoll, Karen M.
APPLICANT: Si, Joan Q.
APPLICANT: Aagaard, Jesper
TITLE OF INVENTION: Glucose Oxidases
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5834280disk of No. 5834280th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,283
FILING DATE: 07-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Agria, Dr. Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 4158,214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 867-0123
TELEFAX: (212) 878-9655
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 613 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-746-283-5

Alignment Scores:
Pred. No.: 43.6 Length: 613
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
Gaps: 0
DB: 2

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-746-283-5 (1-613)

QY 1593 CGTGAGTGCTGCTGCTGTGGA 1570
Db 308 ArgGlvalleuLeuAlaGlyGly 315

RESULT 42
US-08-746-257A-1
Sequence 1, Application US/08746257A
Patent No. 5879921
GENERAL INFORMATION:
APPLICANT: Cherry, Joel
APPLICANT: Berka, Randy
APPLICANT: Halkier, Torben
TITLE OF INVENTION: Glucose Oxidases
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 58799210 No. 5879921disk of No. 5879921th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,257A
FILING DATE: 07-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Starnes, Robert L.
REGISTRATION NUMBER: 41,324
REFERENCE/DOCKET NUMBER: 4700,000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 613 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-746-257A-1

Alignment Scores:
Pred. No.: 43.6 Length: 613
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
Gaps: 0
DB: 2

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-746-257A-1 (1-613)

QY 1593 CGTGAGTGCTGCTGCTGTGGA 1570
Db 308 ArgGlvalleuLeuAlaGlyGly 315

RESULT 43
US-09-347-878-48
Sequence 48, Application US/09347878C
Patent No. 6376210
GENERAL INFORMATION:
APPLICANT: Yuan, Cheng
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
FILE REFERENCE: 25885-1651
CURRENT APPLICATION NUMBER: US/09/347,878C
CURRENT FILING DATE: 1999-07-06
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 48
LENGTH: 613
TYPE: PRT
ORGANISM: Cladosporium oxysporum
FEATURE:
OTHER INFORMATION: Glucose oxidase protein sequence
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 08/746,257
PATENT FILING DATE: 1996-11-07
PUBLICATION DATE: 1999-03-09
US-09-347-878-48

Alignment Scores:
Pred. No.: 43.6 Length: 613
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
Gaps: 0
DB: 4

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-347-878-48 (1-613)

QY 1593 CGTGAGTGCTGCTGCTGTGGA 1570

```
Db      308 ArgGlValLeuLeuLagLygLy 315
|||||
RESULT 44
US-08-472-888A-6
; Sequence 6, Application US/08472888A
; Patent No. 6613746
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Walz, Gerd
; TITLE OF INVENTION: AGP-ANTIBODY FUSION PROTEINS
; TITLE OF INVENTION: AND RELATED MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,888A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/618,314
; FILING DATE: 23-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/258001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TEXT:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-472-888A-6
Alignment Scores:
Pred. No.: 43.4 Length: 630
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
Gaps: 0
DB: 4

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-472-888A-6 (1-630)
QY      1819 GACGAGCTGACTCGCGCGGAGC 1842
|||||
Db      78 AspArgAlaAspSerArgArgSer 85
|||||
RESULT 45
US-09-591-095-6
; Sequence 6, Application US/09591095
; Patent No. 6489461
; GENERAL INFORMATION:
; APPLICANT: Froman, B. & Dehesh, K.
; TITLE OF INVENTION: Nucleic Acid Sequences encoding Proteins
; TITLE OF INVENTION: Involved in Fatty Acid Beta-Oxidation and Methods of Use.
; FILE REFERENCE: 17137/01/US
; CURRENT APPLICATION NUMBER: US/09/591,095
```

```
; CURRENT FILING DATE: 2000-06-08
; EARLIER APPLICATION NUMBER: US 60/138,162
; EARLIER FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 683
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-09-591-095-6
Alignment Scores:
Pred. No.: 43.1 Length: 683
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
Gaps: 0
DB: 4

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-591-095-6 (1-683)
QY      596 AGCTGACTTATCCCGAGTCTC 619
|||||
Db      25 SerLeuThrLeuSerProSerIeu 32
|||||
RESULT 46
US-08-021-601-12
; Sequence 12, Application US/08021601
; Patent No. 5591631
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Nichols, Peter J.
; APPLICANT: Atora, Naveen
; APPLICANT: Singh, Yogendra
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Needle & Rosenberg, P.C.
; STREET: 133 Carnegie Way, Suite 400
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,601
; FILING DATE: 19930212
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-021-601-12
Alignment Scores:
Pred. No.: 41.9 Length: 903
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 1 Gaps: 0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-021-601-12 (1-903)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842

DB 778 AspaRgAlaAspSerArgArgSer 785

RESULT 47

US-08-082-849B-12

Sequence 12, Application US/08082849B

Patent No. 5677274

GENERAL INFORMATION:

APPLICANT: Leppia, Stephen H.

APPLICANT: Klimpel, Kurt R.

APPLICANT: Aroza, Naveen

APPLICANT: Singh, Yogendra

APPLICANT: Nichols, Peter J.

TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and

TITLE OF INVENTION: Related Methods

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/082,849B

FILING DATE: 25-JUN-1993

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/021,601

FILING DATE: 12-FEB-1993

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677

REFERENCE/DOCKET NUMBER: 15280-161-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 903 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-082-849B-12

Alignment Scores:

Pred. No.: 41.9

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.20%

DB: 1

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-08-082-849B-12 (1-903)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842

DB 778 AspaRgAlaAspSerArgArgSer 785

RESULT 48

PCT-US94-01624-12

Sequence 12, Application PC/TUS9401624

GENERAL INFORMATION:

APPLICANT: Leppia, Stephen H.

APPLICANT: Klimpel, Kurt R.

APPLICANT: Aroza, Naveen

APPLICANT: Singh, Yogendra

APPLICANT: Nichols, Peter J.

TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND

TITLE OF INVENTION: RELATED METHODS

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW

STREET: Stewart Street Tower, 20th Floor, One Market

STREET: Plaza

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/01624

FILING DATE: June 25, 1993

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677

REFERENCE/DOCKET NUMBER: 15280-115

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 543-9600

TELEFAX: (415) 543-5043

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 903 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US94-01624-12

Alignment Scores:

Pred. No.: 41.9

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.20%

DB: 5

US-10-029-345A-108_COPY_532_2532 (1-2001) x PCT-US94-01624-12 (1-903)

QY 1819 GACAGAGCTGACTCGCGCGGAGC 1842

DB 778 AspaRgAlaAspSerArgArgSer 785

RESULT 49

US-09-252-991A-33037

Sequence 33037, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NOCLETIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 33037

; LENGTH: 935
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-33037

Alignment Scores:

Pred. No.:	41.7	Length:	935
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-252-991A-33037 (1-935)

QY 1122 GGACAGCCCGCTGTGTACAGCGCT 1145

DB 307 GlyGlnProAlaGlyThrGlyAla 314

RESULT 50

US-09-252-991A-27230
; Sequence 27230, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27230
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27230

Alignment Scores:

Pred. No.:	41.6	Length:	953
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.24%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345A-108_COPY_532_2532 (1-2001) x US-09-252-991A-27230 (1-953)

QY 1107 CTGCACGCTGGGACGCTGGGCAC 1084

DB 241 LeuHisAlaGlyHisAlaGlyHis 248

Search completed: June 21, 2004, 13:15:47
Job time : 62.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 23:40:27 ; Search time 12406 Seconds

(without alignments)
6969.962 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccatgagatgatggtg.....tggaatcatgaggtctcc 1995

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1263658

Minimum DB seq length: 1419

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_hng:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_mu:*
19: em_mu:*
20: em_om:*
21: em_ox:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sbs:*
28: em_un:*
29: em_vt:*
30: em_hng_hum:*
31: em_hng_hum:*
32: em_hng_other:*
33: em_hng_mus:*
34: em_hng_pln:*
35: em_hng_rod:*
36: em_hng_mam:*
37: em_hng_vtr:*
38: em_sy:*
39: em_hngo_hum:*
40: em_hngo_mus:*
41: em_hngo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1995	100.0	5450	6	AX482439	Sequence
2	1995	100.0	5450	6	AX482478	Sequence
3	1893	94.9	1998	6	AX260342	Sequence
4	1893	94.9	3059	6	AX278461	Sequence
5	1893	94.9	3496	6	AX441210	Sequence
6	1893	94.9	3521	9	AB052156	Homo sapi
7	1893	94.9	3544	6	AX260340	Sequence
8	1893	94.9	3566	9	AF506796	Homo sapi
9	1893	94.9	3766	6	AX374994	Sequence
10	1893	94.9	4790	6	BD171157	Novel gen
11	1893	94.9	4790	6	BD183422	Novel gen
12	1893	94.9	4790	6	AB051487	Homo sapi
13	1842	92.3	2732	6	AX180875	Sequence
14	1842	92.3	3104	6	AX405700	Sequence
15	1791	89.8	2102	6	AX713989	Sequence
16	1791	89.8	2102	9	AK055973	Homo sapi
17	1791	89.8	5111	6	AX482372	Sequence
18	1696	85.0	2807	6	BC031643	Sequence
19	1593	79.8	2118	6	AX099933	Sequence
20	1418	71.1	1935	9	AY038927	Homo sapi
21	1418	71.1	3332	6	AX441229	Sequence
22	1284	64.4	3284	9	BC042101	Homo sapi
23	1233	61.8	2071	6	AX921917	Sequence
24	1183	59.3	172206	9	AC007619	Homo sapi
25	1183	59.3	188344	2	AC131617	Homo sapi
26	1182	59.2	2200	6	AX921915	Sequence
27	788	39.5	1916	6	AX835305	Sequence
28	788	39.5	1916	9	AK098310	Homo sapi
29	143	7.2	5047	9	HSMB004384	Homo sapi
30	79	4.0	2756	6	AX482444	Sequence
31	79	4.0	4874	10	AB052157	Mus muscu
32	79	4.0	4943	10	BC059232	Mus muscu
33	79	4.0	4975	10	BC057321	Mus muscu
34	79	4.0	201474	2	AC126692	Mus muscu
35	79	4.0	236589	2	AC118035	Mus muscu
36	76	3.8	244605	2	AC097818	Rattus no
37	76	3.8	250782	2	AC128093	Rattus no
38	72	3.6	3661	10	AF345954	Mus muscu
39	72	3.6	4026	10	AF345953	Mus muscu
40	72	3.6	4827	10	AF345952	Mus muscu
41	72	3.6	4992	10	AF345951	Mus muscu
42	53	2.7	201474	2	AC126692	Mus muscu
43	40	2.0	247148	2	AC097776	Rattus no
44	38	1.9	242590	2	AC133722	Rattus no
45	22	1.1	84077	9	AL157371	Human DNA

ALIGNMENTS

RESULT 1	AX482439	5450 bp	DNA	linear	PAT 16-AUG-2002
LOCUS	AX482439				
DEFINITION	Sequence 108 from Patent WO02057460.				
ACCESSION	AX482439				
VERSION	AX482439.1	GI:22316984			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	Toddard,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G., Suchard,S., Banas,D., Bassolino,D., Feder,J., Kyslek,S., Mcatee,P., Mintier,G., Siemers,N., Jackson,D.G. and Ramamathan,C.				

TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 108 25-JUL-2002;
SOUTHB BRISTOL MYERS CO (US)

FEATURES
Source location/Qualifiers

CDS
1..5450
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
538..2535
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD44451.1"
/db_xref="GI:22316985"
/translation="MAHMGITGVTERVLALESSTKELVLDISRPVYNTSHILE
ALININGSKIMRRLQODKVLITELIOLSHAKYDIDDSQVAVYIDDSQVASSSDC
FLVYLKLEKSRNSVHLHLAGFAEFERCPGLQEKSTLVPTCISQPCLPVANIPT
RIIPNLVIGQCDVLANKEILQONGIYVLANASYCPKDPDI PSHFLRVPVNDSPFK
ILPWLDSVDFIERAKASNGCVLHVCHLAGISRSATLAIAYIMKRMDSIDEAVRYE
KRPTISPNFNLGQLDYKKIKNOTGASGPKSKLHLKEMNEPVPVASEGQKSE
TPLSPCADSATSEBAGORPVHPASVPSVPSVOPLLEDSPLVALSGHLISDRLED
SNLKRSEFSDIKSVSASMAASLHGSSSPALHYKRSTLIDGNLKCPSFVQE
LSBOTPTSPDKPEASIPKYLQTPRPDSQSKLHSVRSSSGTAGRSLSPLHRSQS
VEDNYHTSPFLGSLSTSOQHLTFKAGLKLKGMHSDILAPQSTSLTSSWYFATESSH
YSATATYGSASISAVSCQLPTCGDQVYSVRROKPSRADSRSRWHERSPKQPK
RRCQMEFGESIMSENRSBELKVGSSFSGSMELIEVS"

ORIGIN

Query Match 100.0%; Score 1995; DB 6; Length 5450;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGCCCATGATGATGATGAACTCAATGTTACTGAGAGTGTGCTCTGTGAA 60
DB 538 ATGGCCCATGATGATGATGAACTCAATGTTACTGAGAGTGTGCTCTGTGAA 597
QY 61 AGTGGAAACGAAAAAGTCTGCTAATTGATAGCCGCGCATTTGTGAATACATACCC 120
DB 598 AGTGGAAACGAAAAAGTCTGCTAATTGATAGCCGCGCATTTGTGAATACATACCC 657
QY 121 CACATTTTGGAGCATTAATCACTGCTCCAGCTTATGAAAGGAAGTTGCAACG 180
DB 658 CACATTTTGGAGCATTAATCACTGCTCCAGCTTATGAAAGGAAGTTGCAACG 717
QY 181 GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCAAAACATAGGTGACATGAT 240
DB 718 GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCAAAACATAGGTGACATGAT 777
QY 241 TGCAGTCAGAAAGTTGATGATTCGATCAAAAGTCCCAAGATGTTCCCTCTCTTCA 300
DB 778 TGCAGTCAGAAAGTTGATGATTCGATCAAAAGTCCCAAGATGTTCCCTCTCTTCA 837
QY 301 GACTGTTTCTCACTGACTTCTGAGTAACTGAGAGAAGCTTCAACTCTGTACCGG 360
DB 838 GACTGTTTCTCACTGACTTCTGAGTAACTGAGAGAAGCTTCAACTCTGTACCGG 897
QY 361 CTTCGAGGTGGTTTCTGAGTCTCTGTTGTTCCCTGGGCTCTGTGAAGAAATCC 420
DB 898 CTTCGAGGTGGTTTCTGAGTCTCTGTTGTTCCCTGGGCTCTGTGAAGAAATCC 957
QY 421 ACTCTAGCCCTACCGGATTTCTCAGCCTTGTACTTACCTGTGGCAACATTGGGCAAC 480
DB 958 ACTCTAGCCCTACCGGATTTCTCAGCCTTGTACTTACCTGTGGCAACATTGGGCAAC 1017
QY 481 CGAATTTCTCCAACTTTATCTTGAGTCCAGCAGAGATGCTCTCAACAGAGAGCTGATA 540
DB 1018 CGAATTTCTCCAACTTTATCTTGAGTCCAGCAGAGATGCTCTCAACAGAGAGCTGATA 1077
QY 541 CAGCAGATGGATTTGTTATGTTAAATGCACTATACCTGTCCAAAGCTGACTTT 600
DB 1078 CAGCAGATGGATTTGTTATGTTAAATGCACTATACCTGTCCAAAGCTGACTTT 1137
QY 601 ATCCCGAGTCTCATTTCTGCGTGGCTGTGATGACAGCTTTTGTGAGAAATTTTG 660
```

```
DB 1138 ATCCCGAGTCTCATTTCTGCGTGGCTGTGATGACAGCTTTTGTGAGAAATTTTG 1197
QY 661 CCGTGGTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCCTCCAAATGGATGTT 720
DB 1198 CCGTGGTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCCTCCAAATGGATGTT 1257
QY 721 CTATGTCATGTTTACGTGGGATCTCCGCTCCGCAACCATGCTATGCTTACATG 780
DB 1258 CTATGTCATGTTTACGTGGGATCTCCGCTCCGCAACCATGCTATGCTTACATG 1317
QY 781 AAGAGATGAGATGCTTTAGATGAGCTTACATTTTGAAGAAAGAAAGAAAGCTTACT 840
DB 1318 AAGAGATGAGATGCTTTAGATGAGCTTACATTTTGAAGAAAGAAAGAAAGCTTACT 1377
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGCACTATAGAGAAAGATTAGAAC 900
DB 1378 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGCACTATAGAGAAAGATTAGAAC 1437
QY 901 CAGACTGGAGCTTCAGGGCCAAAGCAAACTCAAGCTGTGCACTTGAGAGCCCAAT 960
DB 1438 CAGACTGGAGCTTCAGGGCCAAAGCAAACTCAAGCTGTGCACTTGAGAGCCCAAT 1497
QY 961 GAACCTGTCCCTGCTGTCTCAGAGGGTGGAGAGAAAGCGAGAGCCCTCACTGCCACC 1020
DB 1498 GAACCTGTCCCTGCTGTCTCAGAGGGTGGAGAGAAAGCGAGAGCCCTCACTGCCACC 1557
QY 1021 TGTGCCACTGTGCTACTCTCAGAGGCAAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1080
DB 1558 TGTGCCACTGTGCTACTCTCAGAGGCAAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1617
QY 1081 CCCAGCGTCCAGCGTGGCAAGCCGCTGTTAAGAGACAGCCCGCTGTGACAGCGCTC 1140
DB 1618 CCCAGCGTCCAGCGTGGCAAGCCGCTGTTAAGAGACAGCCCGCTGTGACAGCGCTC 1677
QY 1141 AGTGGCTGCACTGTCCGACAGAGCTGGAGAGACAAATAGCTCAAGCGTTCCCTTC 1200
DB 1678 AGTGGCTGCACTGTCCGACAGAGCTGGAGAGACAAATAGCTCAAGCGTTCCCTTC 1737
QY 1201 TCTCTGATATCAATCATGTTTCAATATTACGCGAGCATGGCAGCATCTTACATGCGTTC 1260
DB 1738 TCTCTGATATCAATCATGTTTCAATATTACGCGAGCATGGCAGCATCTTACATGCGTTC 1797
QY 1261 TCTCATCAGAAAGTGGCTTGGAAATCTACAAACCTTCCATCTGTGATGGAGCAAC 1320
DB 1798 TCTCATCAGAAAGTGGCTTGGAAATCTACAAACCTTCCATCTGTGATGGAGCAAC 1857
QY 1321 AAGCTATGCACTTCCCTGTTACAGAACTATCGAGAGAGACTCCCGAAACCAAGTCT 1380
DB 1858 AAGCTATGCACTTCCCTGTTACAGAACTATCGAGAGAGACTCCCGAAACCAAGTCT 1917
QY 1381 GATTAAGAGAGAGCAGCATCTCCCAAGAGCTGACAGCCGACGCTTACAGAGCAG 1440
DB 1918 GATTAAGAGAGAGCAGCATCTCCCAAGAGCTGACAGCCGACGCTTACAGAGCAG 1977
QY 1441 AGCAAGGATTTGCAATTTGGTCAAGAACAGCAGAGTGGACCGCCAGAGGTCCCTTTTA 1500
DB 1978 AGCAAGGATTTGCAATTTGGTCAAGAACAGCAGAGTGGACCGCCAGAGGTCCCTTTTA 2037
QY 1501 TCTTCATCTGATGAAAGTGGAGCGTGGAGAGCAATTAACAACAAGCTTCTTTCCGCG 1560
DB 2038 TCTTCATCTGATGAAAGTGGAGCGTGGAGAGCAATTAACAACAAGCTTCTTTCCGCG 2097
QY 1561 CTTTCCACGACGACGACGACCTCAAGAACTGCTGGGCTTGAAGGGCTGGCAC 1620
DB 2098 CTTTCCACGACGACGACGACCTCAAGAACTGCTGGGCTTGAAGGGCTGGCAC 2157
QY 1621 TCGGATATCTTGGCCCCCAGAGCTCTACGCCCTTCCCTGACAGCAGCTGATTTTGGC 1680
DB 2158 TCGGATATCTTGGCCCCCAGAGCTCTACGCCCTTCCCTGACAGCAGCTGATTTTGGC 2217
QY 1681 ACAGAGTCTTACACTTCTGCTGCTGAGCATGAGAGCAGTCCAGTTACTCT 1740
```

[illegible]

LOCUS	AX482478	5450 bp	DNA	linear	PAT 16-AUG-2002
DEFINITION	Sequence 147 from Patent WO202057460.				
ACCESSION	AX482478				
VERSION	AX482478.1	GI:22316999			

SOURCE ORGANISM	Homo sapiens (human) Homo sapiens
-----------------	--------------------------------------

REFERENCE
AUTHORS
1
Toddardud, C.G., Bol, D., Finger, J., Lee, L., Nelson, T., Schieven, G.,

TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 147 25-UTL-2002;
SQUIBB BRISTOL MYERS CO (US)

FEATURES	Location/Qualifiers
source	1. .5450

```

/db_xref="taxon:9606"
538. .2535
CDS

```

```

//sequence_id="2"
//protein_id="CAD44453.1"
/db_xref="GI:22317000"
/db_xref="REMBL:CAD44453"
/translation="MAHEMIGIOIYTEELVALLLESGETKVLIDSRPVEVNTSHILE
AININCSKLMKRLQODKVLITELIOHAKKKKVIDCOKVYVVDOSADVASLSDC
FLTYLGLTEKSPNSVHLIAGSFAEFSCFPGLCGKSTLYPTCTSDCLVANIICPT
RILPMLDYGCDVDYLNKELIQONGIGVYLVANSYTCPEKDFIPESHFLKPVNDVSFCEE
ILPMLDKSVDFIERKANGCVLHVGLAGLIGRAISATIAAYIMKRDMSIDIAFYVFK
KRPTISPMFNFLGOLDIYEKKIKNOTGASKSPKSKLHLKELKPEVPAVSEGGOKS
TLPSPCADSATSEALOGORPVHAPSVSPVOPSLKEDLPVLAOLSLHLSARLEK
SNKIKRPSLIDIKSVYSASMAAHLHGSSSEDLALFYKKPTSDGNKILCOPSPVOS
LSEQTPETSPPDKENASTPKUQTAPASPSIOGSKRLHSVTVSSGTRQORSLLPRLRSG
VEDNHTHSFLFLGLSTSOQHLTKRSAGLILKGMHSDILAQOTSPSLTSSWYATSSHS
YASASIIYGGASVSAYVSCOLPTCGDQVYVYRRRCQEDPRADRSRWMEESPFEKQF
RRSCOMEGESIMSENREELGKGVSSSSFSGSEILIEVA"

```

ORIGIN

Query	Match Similarity	Score	1995:	DB	6:	Length	5450:
Beef	Local	Similarity	100.0%	Pred.	No.	0:	
Matches	1995:	Conservative	0:	Mismatches	0:	Indels	Gaps
Qy	1	ATGCGCCATGAGATATTGGAACTCAATTTGTACTGAGAGGTTGTCGCTCTGGGAA	60				
Db	538	ATGCGCCATGAGATATTGGAACTCAATTTGTACTGAGAGGTTGTCGCTCTGGGAA	597				

Oy	61	GGTGAACGGAAAAAGCGCTGTAATTGATGCGCGGCATTGTGGAAATACAATCATCC	120
Db	598	AGTGAACGGAAAAAGCGCTGCTAATTGATGCGCGGCATTGTGGAAATACAATCATCC	657
Oy	121	CACATTTTGGAAAGCCATTAAATATCAACTGCTCCAAAGCTTATGAAAGCAAGTTGGCAACG	180
Db	658	CACATTTTGGAAAGCCATTAAATATCAACTGCTCCAAAGCTTATGAAAGCAAGTTGGCAACG	717
Oy	181	GACAAAGTGTAAATTAACAGAGCTCATCCAGATTCAGCGAAACATTAAGTTGACATTGAT	240
Db	718	GACAAAGTGTAAATTAACAGAGCTCATCCAGATTCAGCGAAACATTAAGTTGACATTGAT	777
Oy	241	TGCAGTCAGAAGGTTGTAGTTAAGATPAAAGCTCCCAAGATGTGCTCTCTCTTCA	300
Db	778	TGCAGTCAGAAGGTTGTAGTTAAGATPAAAGCTCCCAAGATGTGCTCTCTCTTCA	837
Oy	301	GACGTTTTCACACTGACTTCTGGGGTAACTGGAGAAGAGCTTCAACTGTGTACCTG	360
Db	838	GACGTTTTCACACTGACTTCTGGGGTAACTGGAGAAGAGCTTCAACTGTGTACCTG	897

QY	421	ACTGATGCCCTACTGATCTTCTCAGGCTTGCTTACCTGTTGGCAACATTTGGGCAAC	480
Db	958	ACTGATGCCCTACTGATCTTCTCAGGCTTGCTTACCTGTTGGCAACATTTGGGCAAC	101
QY	481	CGAATCTTCCCAATCTTTATCTTGCGTCCAGCGAGATGTCTTCMAAGAGAGCTGATA	540
Db	1018	CGAATCTTCCCAATCTTTATCTTGCGTCCAGCGAGATGTCTTCMAAGAGAGCTGATA	107
QY	541	CAGCAGAATGGGATTTGGTTATGTGTTAAATGCGACGCTATACCTGTTCMAAGCTTGACTTT	600
Db	1078	CAGCAGAATGGGATTTGGTTATGTGTTAAATGCGACGCTATACCTGTTCMAAGCTTGACTTT	113
QY	601	ATCCCCGAGTCATCTTCTCGCGTGCGCTGTGAAATGACAGCTTTGTGAGAAAATTTTG	660
Db	1138	ATCCCCGAGTCATCTTCTCGCGTGCGCTGTGAAATGACAGCTTTGTGAGAAAATTTTG	119
QY	661	CCGTGTTGACAAATCAGTAGATTTCTATGAGAAAAGCAAAAGCCTCCAAATGATGTGTT	720
Db	1198	CCGTGTTGACAAATCAGTAGATTTCTATGAGAAAAGCAAAAGCCTCCAAATGATGTGTT	125
QY	721	CTAATGCATCTGTTTAGCTGGGATCTCCCGCTCCGCACCAATCGCTATAGGCTTACATCATG	780
Db	1258	CTAATGCATCTGTTTAGCTGGGATCTCCCGCTCCGCACCAATCGCTATAGGCTTACATCATG	131
QY	781	AAGAGGATGACATGTCTTTAGATGAAAGCTTAACAGATTTGTGAAAGAAAAGACCTACT	840
Db	1318	AAGAGGATGACATGTCTTTAGATGAAAGCTTAACAGATTTGTGAAAGAAAAGACCTACT	137
QY	841	ATATCTCCAAACTTCAATTTTTCTGGGCCCAACTCTCTGAATAAGAGAGAGATTAAGAAC	900
Db	1378	ATATCTCCAAACTTCAATTTTTCTGGGCCCAACTCTCTGAATAAGAGAGAGATTAAGAAC	143
QY	901	CAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCAACTCTGAGAAAGCCAAAT	960
Db	1438	CAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCAACTCTGAGAAAGCCAAAT	149
QY	961	GAACCTGTCCCTGCTGTCTCAGAGGGGTGAGACAGAAAACGAGAACGCCCTCAGTCCACCC	1020
Db	1498	GAACCTGTCCCTGCTGTCTCAGAGGGGTGAGACAGAAAACGAGAACGCCCTCAGTCCACCC	155
QY	1021	TGTGCGCACTCTGCTACTCAGAGGCGACAGAGCAAAAGCCGTGTCAATCCGCGACGCTG	1080
Db	1558	TGTGCGCACTCTGCTACTCAGAGGCGACAGAGCAAAAGCCGTGTCAATCCGCGACGCTG	161
QY	1081	CCGAGCGTGCCAGCGTGCAGCCGTCGTGTTAGAGGACAGCCGCTGTGTACAGGCGCTC	1140
Db	1618	CCGAGCGTGCCAGCGTGCAGCCGTCGTGTTAGAGGACAGCCGCTGTGTACAGGCGCTC	167

QY 1141 AGTGGGCTGCACCTGTCCGACAGAGGCTGGAAGACAGCAATPAAGTCAAGGCTTCTTC 1200
| | | | |
Db 1678 AGTGGGCTGCACCTGTCCGACAGAGGCTGGAAGACAGCAATPAAGTCAAGGCTTCTTC 1737
| | | | |
QY 1201 TCTCTGGAATTCAAATCACTTTTCACTTACGACAGATGGGCAATCCCTTACATGGCTTC 1260
| | | | |
Db 1738 TCTCTGGAATTCAAATCACTTTTCACTTACGACAGATGGGCAATCCCTTACATGGCTTC 1797
| | | | |
QY 1261 TCTCTGGAATTCAAATCACTTTTCACTTACGACAGATGGGCAATCCCTTACATGGCTTC 1320
| | | | |
Db 1798 TCTCTGGAATTCAAATCACTTTTCACTTACGACAGATGGGCAATCCCTTACATGGCTTC 1857
| | | | |
QY 1321 AAGCTATGCCAGTTCTTCCCTGTTCGAGAACTATCGGAGCAGACTCCGAAACCAAGTCTT 1380
| | | | |
Db 1858 AAGCTATGCCAGTTCTTCCCTGTTCGAGAACTATCGGAGCAGACTCCGAAACCAAGTCTT 1917
| | | | |
QY 1381 GATPAAGAGGAGAGCAGCAATCCGACAGAGCTGACAGCCGAGGCTTCAAGCAGCCAG 1440
| | | | |
Db 1918 GATPAAGAGGAGAGCAGCAATCCGACAGAGCTGACAGCCGAGGCTTCAAGCAGCCAG 1977
| | | | |
QY 1441 AGCAGAGCAGATTGATTCGATCAGAACAGCAGAGCAGAGCCGACAGAGGCTTCTTTTA 1500
| | | | |
Db 1978 AGCAGAGCAGATTGATTCGATCAGAACAGCAGAGCAGAGCCGACAGAGGCTTCTTTTA 2037
| | | | |
QY 1501 TCTTCACTGATCGAAGTGGGAGCCTGGAGAGCAATTAACAACAAGCTTCTTCTTCCGCG 1560
| | | | |
Db 2038 TCTTCACTGATCGAAGTGGGAGCCTGGAGAGCAATTAACAACAAGCTTCTTCTTCCGCG 2097
| | | | |
QY 1561 CTTTCCACAGCAGCAGCAGCAGCTCAGAGAGTGTGCTGAGCCTTAAAGGCTGGCAG 1620
| | | | |
Db 2098 CTTTCCACAGCAGCAGCAGCAGCTCAGAGAGTGTGCTGAGCCTTAAAGGCTGGCAG 2157
| | | | |
QY 1621 TCGGATATCTTGGCCCCCAGACCTTCACTTCCCTGACAGCAGCTGTATTTTGGC 1680
| | | | |
Db 2158 TCGGATATCTTGGCCCCCAGACCTTCACTTCCCTGACAGCAGCTGTATTTTGGC 2217
| | | | |
QY 1681 ACAGAGCTCAGACCTTCACTTCTGAGCTGAGCCATCAGAGAGCAGAGTGCAGTACTCT 1740
| | | | |
Db 2218 ACAGAGCTCAGACCTTCACTTCTGAGCTGAGCCATCAGAGAGCAGAGTGCAGTACTCT 2277
| | | | |
QY 1741 GCCTACAGCTGACAGCAGCTGACCTTGGGAGAGCAGCAATCTTCTGTGGCAGGCG 1800
| | | | |
Db 2278 GCCTACAGCTGACAGCAGCTGACCTTGGGAGAGCAGCAATCTTCTGTGGCAGGCG 2337
| | | | |
QY 1801 CAGAGAGCAGAGTGAAGAGCTGAGTGGGCGAGAGCTGAGTGAAGAGAGCCCTTTTGA 1860
| | | | |
Db 2338 CAGAGAGCAGAGTGAAGAGCTGAGTGGGCGAGAGCTGAGTGAAGAGAGCCCTTTTGA 2397
| | | | |
QY 1861 AAGCAGTTTAAAGCAGAGCTGCAATATGAAATTTGAGAGAGCATCATGTCAAGAAC 1920
| | | | |
Db 2398 AAGCAGTTTAAAGCAGAGCTGCAATATGAAATTTGAGAGAGCATCATGTCAAGAAC 2457
| | | | |
QY 1921 AAGTCACGGGAGAGAGTGGGAGAAAGTGGGAGTCACTTCTTCTTGGGAGCAGATGAA 1980
| | | | |
Db 2458 AAGTCACGGGAGAGAGTGGGAGAAAGTGGGAGTCACTTCTTCTTGGGAGCAGATGAA 2517
| | | | |
QY 1981 ATCTATTAGAGTCTCC 1995
| | | | |
Db 2518 ATCTATTAGAGTCTCC 2532
| | | | |
RESULT 3
LOCUS AX260342 1998 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 3 from Patent WO0173059.
ACCESSION AX260342
VERSION AX260342.1 GI:16509305
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1

AUTHORS Meyers, R.A.
TITLE 38692 and 21117: dual specificity phosphatase molecules and uses
JOURNAL Patent: WO 0173059-A 3 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source 1..1998
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 94.9%; Score 1893; DB 6; Length 1998;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATTTGAACTCAATTGTTACTGAGAGGTTGGCTCTGCTGGA 60
| | | | |
Db 1 ATGGCCCATGAGATGATTTGAACTCAAAATGTTACTGAGAGGTTGGCTCTGCTGGA 60
| | | | |
QY 61 AGTGAAGGAAAAAGTCTGCTAATGATGCGGCGCATTTGGAAATACATATATCC 120
| | | | |
Db 61 AGTGAAGGAAAAAGTCTGCTAATGATGCGGCGCATTTGGAAATACATATATATCC 120
| | | | |
QY 121 CACATTTTGAAGCATTATATATCACTGCTCAAGCTTATGAAGCGAAGTTGCAAG 180
| | | | |
Db 121 CACATTTTGAAGCATTATATATCACTGCTCAAGCTTATGAAGCGAAGTTGCAAG 180
| | | | |
QY 181 GACAAAGTTAATTACAGAGCTCATCCAGATTCAGGAAACATTAAGTTGACATTGAT 240
| | | | |
Db 181 GACAAAGTTAATTACAGAGCTCATCCAGATTCAGGAAACATTAAGTTGACATTGAT 240
| | | | |
QY 241 TGCAGTCAAGAGGTGTGTTTACATGATCAAGAGTCCCAAGATGTGCTCTCTCTCA 300
| | | | |
Db 241 TGCAGTCAAGAGGTGTGTTTACATGATCAAGAGTCCCAAGATGTGCTCTCTCTCA 300
| | | | |
QY 301 GACTGTTTCTCACTGTAATTTCTGGGTAACCTGAGAAAGCTTCACTCTGTTCACTG 360
| | | | |
Db 301 GACTGTTTCTCACTGTAATTTCTGGGTAACCTGAGAAAGCTTCACTCTGTTCACTG 360
| | | | |
QY 361 CTTCAGAGTGGGTTTGTGAGATTCCTGTTGTTTCCCTGGGCTCTGTAAGGAAATATCC 420
| | | | |
Db 361 CTTCAGAGTGGGTTTGTGAGATTCCTGTTGTTTCCCTGGGCTCTGTAAGGAAATATCC 420
| | | | |
QY 421 ACTTAGTCCCTACCTGATTTCTCAGGCTTGTCTTACCTGTTGCCAATTGGGCAACC 480
| | | | |
Db 421 ACTTAGTCCCTACCTGATTTCTCAGGCTTGTCTTACCTGTTGCCAATTGGGCAACC 480
| | | | |
QY 481 CGAATTTCTCCCAATCTTATCTTGGCTGCGCAGAGATGTCCTCAACAAGAGCTGATG 540
| | | | |
Db 481 CGAATTTCTCCCAATCTTATCTTGGCTGCGCAGAGATGTCCTCAACAAGAGCTGATG 540
| | | | |
QY 541 CAGCAGATGGGATTTGTTATGTTAATATGCACTTACCTGTCGAAGGCTGACTT 600
| | | | |
Db 541 CAGCAGATGGGATTTGTTATGTTAATATGCACTTACCTGTCGAAGGCTGACTT 600
| | | | |
QY 601 ATCCCGAGTCTCATTTCTCGCTGCTGCTGTAATGAAGCTTTTGTGAAGAAATTTTG 660
| | | | |
Db 601 ATCCCGAGTCTCATTTCTCGCTGCTGCTGTAATGAAGCTTTTGTGAAGAAATTTTG 660
| | | | |
QY 661 CCGTGGTTGACAAATCAGTATGATTTCACTGGAAGAAAGAGCTTCCAAATGAGATGTT 720
| | | | |
Db 661 CCGTGGTTGACAAATCAGTATGATTTCACTGGAAGAAAGAGCTTCCAAATGAGATGTT 720
| | | | |
QY 721 CTAGTGCATGTTTATGCTGGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 780
| | | | |
Db 721 CTAGTGCATGTTTATGCTGGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 780
| | | | |
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
| | | | |
Db 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
| | | | |
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTGAGTATGAGAAAGATTAAGAAC 900
| | | | |

Db 841 ATATCTCAAACTTCAATTTCTGGGCAACTCTGAGCTATGAGAAGAAATTATGAAC 900
Qy 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAGCTGTGCACTCGGAAGCCAAAT 960
Db 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAGCTGTGCACTCGGAAGCCAAAT 960
Qy 961 GAACCTGTCCCTGCTGCTCAGAGGGTGA CAGAAAAGGAGAGCCGCTCAGTCCACCC 1020
Db 961 GAACCTGTCCCTGCTGCTCAGAGGGTGA CAGAAAAGGAGAGCCGCTCAGTCCACCC 1020
Qy 1021 TGTGCCGACTCTGCTACTCCTCAGAGGAGCAGAGCAAAAGCCGCTGATCCCGCAGCGTG 1080
Db 1021 TGTGCCGACTCTGCTACTCCTCAGAGGAGCAGAGCAAAAGCCGCTGATCCCGCAGCGTG 1080
Qy 1081 CCCAGGCTGCCAGCGCTGACGCGCTGCTTTAGAGAGACGCCCCGCTGGTACAGGCGCTC 1140
Db 1081 CCCAGGCTGCCAGCGCTGACGCGCTGCTTTAGAGAGACGCCCCGCTGGTACAGGCGCTC 1140
Qy 1141 AGTGGGCTGCACTGTCCGAGCAGGGCTGGAAGACGAACTAAGCTCAAGCGTCTTC 1200
Db 1141 AGTGGGCTGCACTGTCCGAGCAGGGCTGGAAGACGAACTAAGCTCAAGCGTCTTC 1200
Qy 1201 TCTCTGGAATCAAACTCAGTTTATATTCAGCCAGATGAGCATCTTACATGAGCTTC 1260
Db 1201 TCTCTGGAATCAAACTCAGTTTATATTCAGCCAGATGAGCATCTTACATGAGCTTC 1260
Qy 1261 TCTCTCAGAGAGATGCTTTGGAATACAACTTCCACTCTGATGGAGCAAC 1320
Db 1261 TCTCTCAGAGAGATGCTTTGGAATACAACTTCCACTCTGATGGAGCAAC 1320
Qy 1321 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATCGAGAGACTCCGAAACCACTCTC 1380
Db 1321 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATCGAGAGACTCCGAAACCACTCTC 1380
Qy 1381 GATTAAGAGAGAGCCAGCATCCCGCAAGAGCTCAGACCGCCCTTACAGACGAC 1440
Db 1381 GATTAAGAGAGAGCCAGCATCCCGCAAGAGCTCAGACCGCCCTTACAGACGAC 1440
Qy 1441 AGCAGCGATTCATTCGCTCAGAACCAAGCAGAGTGGCACCGCCCAAGCTCTTTTA 1500
Db 1441 AGCAGCGATTCATTCGCTCAGAACCAAGCAGAGTGGCACCGCCCAAGCTCTTTTA 1500
Qy 1501 TCTCCACTGATGAGAGTGGAGAGCTGAGAGCAATTAACAACGACTCTTTTCCG 1560
Db 1501 TCTCCACTGATGAGAGTGGAGAGCTGAGAGCAATTAACAACGACTCTTTTCCG 1560
Qy 1561 CTTTCCACGAGCAGAGCACTCAGAAAGTCTGCTGAGCTTGAAGGCTGGCAC 1620
Db 1561 CTTTCCACGAGCAGAGCACTCAGAAAGTCTGCTGAGCTTGAAGGCTGGCAC 1620
Qy 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGACCTGTATTTTGGC 1680
Db 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGACCTGTATTTTGGC 1680
Qy 1681 ACGAGTCTCTCAACTTTACTCTGCTCAGGCACTTACGAGAGGCACTGCAATTTACT 1740
Db 1681 ACGAGTCTCTCAACTTTACTCTGCTCAGGCACTTACGAGAGGCACTGCAATTTACT 1740
Qy 1741 GCTTACAGCTGAGCAGCTGAGCACTTGGAGAGCAAGTCTATTCTGAGCAGAGCG 1800
Db 1741 GCTTACAGCTGAGCAGCTGAGCACTTGGAGAGCAAGTCTATTCTGAGCAGAGCG 1800
Qy 1801 CAGAGCCAAAGTACAGAGCTGACTCGCGGCGAGCTGCGATGAGAGAGCCCTTTTGA 1860
Db 1801 CAGAGCCAAAGTACAGAGCTGACTCGCGGCGAGCTGCGATGAGAGAGCCCTTTTGA 1860
Qy 1861 AAGCAGTTTAAAGCAGAGAGCTGCAATATGAAATTTGAGAGAGCATCATGTACAGAAC 1920
Db 1861 AAGCAGTTTAAAGCAGAGAGCTGCAATATGAAATTTGAGAGAGCATCATGTACAGAAC 1920
Qy 1921 AGGTCAAGGAGAGAGCTGGGGAAGTGGGAGTCACTGCTTTTCCGAGCAGCATGGA 1980
Db 1921 AGGTCAAGGAGAGAGCTGGGGAAGTGGGAGTCACTGCTTTTCCGAGCAGCATGGA 1980

Db 1921 AGGTCAAGGAGAGAGCTGGGGAAGTGGGAGTCACTGCTTTTCCGAGCAGCATGGA 1980
Qy 1981 ATCATTGAGTCTCC 1995
Db 1981 ATCATTGAGTCTCC 1995
RESULT 4
AX278461
LOCUS AX278461 3059 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 1 from Patent WO0177340.
ACCESSION AX278461
VERSION AX278461.1 GI:16605915
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Ducker, K.
AUTHORS Identification of a dual specificity phosphatase: dusp-10
TITLE Patent: WO 0177340-A 1 18-OCT-2001;
JOURNAL MERCK PATENT GmbH (DE)
FEATURES
source 1..3059
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
127..2124
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10545.1"
/db_xref="GI:1605916"
/db_xref="REFSEQ:CAD10545"
/translation="MAHEWIGTQIYTERVALLSEGTREKVLIDSRPPEVNTSHILE
AININCSKLMKRLQODKVLITELIHSKHKVDIDSCQVAVYQDSQDVASLSDC
PLTVLGLKESFNSVHLGAGFAEISRFPELGGKSLVPTCTISQPLCVANTGPT
RLPLNLYIGCCQDVANKELMQONGCYUNASNTCPKPFIESHLRPVNDPEK
ILPWLDSKDFLEKASNGCVLVHLCGISATTAAYIKRMDSIDEAYRFYKE
KRPTISPNENFLGLDYEKLIKNGTASGPYSKULHLLEKPNBPVAVSGGQSE
TPLSPPCDASATSEBAQRPVHPASVPVSPVSLIEDSPVLAHLSGLHLSADRLD
SKLRSFELDIKSVSYASMAASLHGFSSSDALEYKPSYTLDTGNTLCQFSPVQ
LSEOTPERTSPDKESASIPKLIOTARPDSGSRHSVRSSTGQORSLPLSHRGS
VDNNTSTFLRELSTSOCHLTYSAGLGLKGMHSDILAPOTSPSLTSMYPRATESHP
YSASLYGSASISVYSCQPLFCGQVYSVRKQPSRQPSRDSKRSMBESPEKQFK
RRSCMEGFESIMSENREBELKGVQSQSFSGSMELIEVS"

ORIGIN
Query Match 94.9%; Score 1893; DB 6; Length 3059;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGCCCATGAGATGATGGAATCTCAAAATTTGTAAGAGGTGGCTGTCTGGA 60
Db 127 ATGCCCATGAGATGATGGAATCTCAAAATTTGTAAGAGGTGGCTGTCTGGA 186
Qy 61 AGTGAACGGAAGAAAGTGTCTTAATTGATAGCCGGCCATTGTGGAATACATACCTC 120
Db 187 AGTGAACGGAAGAAAGTGTCTTAATTGATAGCCGGCCATTGTGGAATACATACCTC 246
Qy 121 CACATTTTGAAGCCATTAAATATCACTGCTCCAGCTTATAGAGGAAGTTGCAACAG 180
Db 247 CACATTTTGAAGCCATTAAATATCACTGCTCCAGCTTATAGAGGAAGTTGCAACAG 306
Qy 181 GACAAAGTTTAATTAACAGAGCTCATCCAGATTCAGGAAACATTAAGTTGACATTGAT 240
Db 307 GACAAAGTTTAATTAACAGAGCTCATCCAGATTCAGGAAACATTAAGTTGACATTGAT 366
Qy 241 TGCAGTCAGAGGTTGTAGTTTAAGATCAACAGCTCCCAAGATGTGCTCTCTTCA 300
Db 367 TGCAGTCAGAGGTTGTAGTTTAAGATCAACAGCTCCCAAGATGTGCTCTCTTCA 426
Qy 301 GACTGTTTCTCACTGTACTTCTGGGTAAACTGAGAGAGCTTCAACTCTGTACCTG 360

```

Db      427 GACTGTTTCTCACTGTACTTCTGGGTAACGTGAGAAAGCTTCAACTCTGTTCACTG 486
QY      361 CTTGCAAGTGGGTTTGGTGAAGTTCTCTGTTGTTTCCCTGCGCTCTGTGAAGAAATCC 420
Db      487 CTTGCAAGTGGGTTTGGTGAAGTTCTCTGTTGTTTCCCTGCGCTCTGTGAAGAAATCC 546
QY      421 ACTCTAGTCCCTACCTGCAATTTCTAGCCCTTGTCTACCTGTTGCTGCAACATTTGGGCAACC 480
Db      547 ACTTAGTCCCTACCTGCAATTTCTAGCCCTTGTCTACCTGTTGCTGCAACATTTGGGCAACC 606
QY      481 CGAATTTCTCCCAATCTTTATCTTTGCTGCGACAGATGTCTCTCAACAGAGCTGATA 540
Db      607 CGAATTTCTCCCAATCTTTATCTTTGCTGCGACAGATGTCTCTCAACAGAGCTGATA 666
QY      541 CAGCAGATGGAGTTGGTTATGTGTTAAATGCGAGCTATTCCTGTCCAAAGCCTGACTT 600
Db      667 CAGCAGATGGAGTTGGTTATGTGTTAAATGCGAGCAATCTGTCCAAAGCCTGACTT 726
QY      601 ATCCCGAGTCTGATTTCCGCGTGGCTGTGATGACAGCTTTGTGGAATAATTTTG 660
Db      727 ATCCCGAGTCTGATTTCCGCGTGGCTGTGATGACAGCTTTGTGGAATAATTTTG 786
QY      661 CCGTGGTTGCAAAATCAGTATGATTTCAATTGAGAAAACAAGCCTCCAAATGATGTT 720
Db      787 CCGTGGTTGCAAAATCAGTATGATTTCAATTGAGAAAACAAGCCTCCAAATGATGTT 846
QY      721 CTAGTGCATCTTTAGCTGGGATCTCCCGTCCGCCACCTGCTATTCGCTATCATCATG 780
Db      847 CTAGTGCATCTTTAGCTGGGATCTCCCGTCCGCCACCTGCTATTCGCTATCATCATG 906
QY      781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 840
Db      907 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 966
QY      841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAAGAC 900
Db      967 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAAGAC 1026
QY      901 CAGACTGGAGCATCAGAGGCGCAAGAGCAAACTCAAGCTGTGACCTTGAGAAAGCCAAAT 960
Db      1027 CAGACTGGAGCATCAGAGGCGCAAGAGCAAACTCAAGCTGTGACCTTGAGAAAGCCAAAT 1086
QY      961 GAACCTGTCCCTGCTCTCAGAGGTTGACAGAAAAGCAGAGCCCTCACTGCCACC 1020
Db      1087 GAACCTGTCCCTGCTCTCAGAGGTTGACAGAAAAGCAGAGCCCTCACTGCCACC 1146
QY      1021 TGTGCCGACTCTGCTACTCAGAGGACAGAGCAAAAGCCGTGATCTCCGACAGCTG 1080
Db      1147 TGTGCCGACTCTGCTACTCAGAGGACAGAGCAAAAGCCGTGATCTCCGACAGCTG 1206
QY      1081 CCCAGGTCGCCGACGCTGACGCGTGTCTTTAGAGACAGCCCGCTGTGATCAGGGCTC 1140
Db      1207 CCCAGGTCGCCGACGCTGACGCGTGTCTTTAGAGACAGCCCGCTGTGATCAGGGCTC 1266
QY      1141 AGTGGCTGCACTGTCTCCGACAGAGCTGGAAGACAGCATATAGCTCAAGCTTCTTC 1200
Db      1267 AGTGGCTGCACTGTCTCCGACAGAGCTGGAAGACAGCATATAGCTCAAGCTTCTTC 1326
QY      1201 TCTCTGATATCAAAATCAAGTTCAATATTCAGCCAGCATGGCAGATCTTACATGGCTTC 1260
Db      1327 TCTCTGATATCAAAATCAAGTTCAATATTCAGCCAGCATGGCAGATCTTACATGGCTTC 1386
QY      1261 TCTCTATCAGAAATAGCTTTGGAATCTCAAACTTCACTATCTCTGATGGAGCAAC 1320
Db      1387 TCTCTATCAGAAATAGCTTTGGAATCTCAAACTTCACTATCTCTGATGGAGCAAC 1446
QY      1321 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATCGAGAGAGCTCCGAAACCACTGCT 1380
Db      1447 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATCGAGAGAGCTCCGAAACCACTGCT 1506
QY      1381 GATTAAGAGAGACGAGCATCCCAAGAAAGCTGACAGCCGCAAGGCTTTCAGACAGCCAG 1440

```

```

Db      1507 GATTAAGAGAGACCGAGCATCCCAAGAGCTGACAGCCGCCAGGCTTTCAGACAGCCAG 1566
QY      1441 AGCAAGCCATTTGCATTTCCGTTCAGAAACGACACACTGGACCGCCAGAGTCCCTTTTA 1500
Db      1567 AGCAAGCCATTTGCATTTCCGTTCAGAAACGACACACTGGACCGCCAGAGTCCCTTTTA 1626
QY      1501 TCTCCACTGCATCAGAGTGGAGCGCTGGAGACAATTAACCAACAGCTTCTCTTTTCGCG 1560
Db      1627 TCTCCACTGCATCAGAGTGGAGCGCTGGAGACAATTAACCAACAGCTTCTCTTTTCGCG 1686
QY      1561 CTTTCCACAGCCAGCAGACACTCAAGATGTGCTGCGCTTGAAGGCTGGCAC 1620
Db      1687 CTTTCCACAGCCAGCAGACACTCAAGATGTGCTGCGCTTGAAGGCTGGCAC 1746
QY      1621 TGGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACACAGAGCTGTATTTTGGC 1680
Db      1747 TGGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACACAGAGCTGTATTTTGGC 1806
QY      1681 ACAGAGTCCCTCACATTTCTACTGCTCCTCAGCATCTACGAGGAGCTGCCAGTTACTCT 1740
Db      1807 ACAGAGTCCCTCACATTTCTACTGCTCCTCAGCATCTACGAGGAGCTGCCAGTTACTCT 1866
QY      1741 GCTTACAGCTGACGCGACGCTGCCACTTGGGAGACCAAGTCTATTTGTGGCAGGCGG 1800
Db      1867 GCTTACAGCTGACGCGACGCTGCCACTTGGGAGACCAAGTCTATTTGTGGCAGGCGG 1926
QY      1801 CAGAAAGCCAAATGACAGAGCTGACTCGCGGCGAGAGCTGGCATGAGAGAGCCCTTTGAA 1860
Db      1927 CAGAAAGCCAAATGACAGAGCTGACTCGCGGCGAGAGCTGGCATGAGAGAGCCCTTTGAA 1986
QY      1861 AAGCATTTTAAACGAGAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCAAGAAC 1920
Db      1987 AAGCATTTTAAACGAGAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCAAGAAC 2046
QY      1921 AGGTACCGGAAAGAGCTGGGAAAGTGGGAGTCACTGCTTTTCCGAGCATGAA 1980
Db      2047 AGGTACCGGAAAGAGCTGGGAAAGTGGGAGTCACTGCTTTTCCGAGCATGAA 2106
QY      1981 ATCATTTGAGGTCTCC 1995
Db      2107 ATCATTTGAGGTCTCC 2121

RESULT 5
AX441210 3496 bp DNA linear PART 26-JUN-2002
LOCUS AX441210
DEFINITION Sequence 1 from Patent WO0226997.
ACCESSION AX441210
VERSION AX441210.1 GI:21665766
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1
AUTHORS Luche, R.M. and Wei, B.
TITLE Dsp-16 dual-specificity phosphatase
JOURNAL Patent: WO 0226997-A 1 04-APR-2002;
Cepcyr, Inc. (US)
FEATURES
source
1..3496
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 94.9%; Score 1893; DB 6; Length 3496;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATTGGAATCAAAATTTGTTACTGAGAGGTTGGTGGCTTGTCTGAA 60
Db 562 ATGGCCCATGAGATGATTGGAATCAAAATTTGTTACTGAGAGGTTGGTGGCTTGTCTGAA 621

```

OY	61	AGTGGAAACGAAAAAAGTCTGCTGATATTTGATTTGCGCGGACATTTGGGAATTAACAATACAC	120
DB	622	AGTGGAAACGAAAAAAGTCTGCTGATATTTGATTTGCGCGGACATTTGGGAATTAACAATACAC	681
OY	121	CACATTTTGGAAAGCCATTAATATATCAACTGCTTCCAGCTTATGAACGAAAGTTGCACAG	180
DB	682	CACATTTTGGAAAGCCATTAATATATCAACTGCTTCCAGCTTATGAACGAAAGTTGCACAG	741
OY	181	GACAAAGTGTTAATTAACAGAGCTCATCCAGCATTTCCAGGAAACATTAAGTTGACATTGAT	240
DB	742	GACAAAGTGTTAATTAACAGAGCTCATCCAGCATTTCCAGGAAACATTAAGTTGACATTGAT	801
OY	241	TGCAGTCAAGAAAGTTGATGATTAACATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA	300
DB	802	TGCAGTCAAGAAAGTTGATGATTAACATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA	861
OY	301	GACTGTTTCTCACTGTACTTCTGGGTAACCTGAGAAAGACTTCAACTCTGTTTCACTG	360
DB	862	GACTGTTTCTCACTGTACTTCTGGGTAACCTGAGAAAGACTTCAACTCTGTTTCACTG	921
OY	361	CTTGGAGGTTGGGTTTGTGAGTTCCTGCTGCTGTTTCCCTGGCCTCTGGAAGGAAATCC	420
DB	922	CTTGGAGGTTGGGTTTGTGAGTTCCTGCTGCTGTTTCCCTGGCCTCTGGAAGGAAATCC	981
OY	421	ACTGTAGTCCCTGACCTGCAATTTCTCAGACCTTGCTTACCTGTTGGCCAACTTGGGCAAC	480
DB	982	ACTGTAGTCCCTGACCTGCAATTTCTCAGACCTTGCTTACCTGTTGGCCAACTTGGGCAAC	1041
OY	481	CGAATTTCTTCCCAATCTTTATCTTGGTGGCTGCCAGAGCAATGTCTCAACAAGAGCTGATA	540
DB	1042	CGAATTTCTTCCCAATCTTTATCTTGGTGGCTGCCAGAGCAATGTCTCAACAAGAGCTGATA	1101
OY	541	CAGCAGATGAGGATGTTGTTATGTGTTAATGCGACAGCTATACCTGTCCAAACCTGACTTT	600
DB	1102	CAGCAGATGAGGATGTTGTTATGTGTTAATGCGACAGCTATACCTGTGTCCAAACCTGACTTT	1161
OY	601	ATCCCGAGTCTCATTTCTCTGGCTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG	660
DB	1162	ATCCCGAGTCTCATTTCTCTGGCTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG	1221
OY	661	CCGTGGTGGACAAATCAGTATGATTCAATTGAGAAAGAAAGAGCTCCAAATGAGATGTGT	720
DB	1222	CCGTGGTGGACAAATCAGTATGATTCAATTGAGAAAGAAAGAGCTCCAAATGAGATGTGT	1281
OY	721	CTAGTGCATGTTTAGCTGGGATCTCCCGCTCCGCAACATCGCTATCGCTTACATCATG	780
DB	1282	CTAGTGCATGTTTAGCTGGGATCTCCCGCTCCGCAACATCGCTATCGCTTACATCATG	1341
OY	781	AAGAGGATGAGACATGCTTTTATGATGATGACCTTACAGATTTGTGAAAGAAAGAAAGCTTACT	840
DB	1342	AAGAGGATGAGACATGCTTTTATGATGATGACCTTACAGATTTGTGAAAGAAAGAAAGCTTACT	1401
OY	841	ATATCTCCAAACTTCAATTTTTCTGGGGCAACCTCGTGAATGAGAAAGAAAGTTAAGAAC	900
DB	1402	ATATCTCCAAACTTCAATTTTTCTGGGGCAACCTCGTGAATGAGAAAGAAAGTTAAGAAC	1461
OY	901	CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGTGAGAAAGCCAAAT	960
DB	1462	CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGTGAGAAAGCCAAAT	1521
OY	961	GAACTGTGCCCTGCTGCTCAAGAGGTTGACAGAAABCGAGAGCCCTCAATGTCACACC	1020
DB	1522	GAACTGTGCCCTGCTGCTCAAGAGGTTGACAGAAABCGAGAGCCCTCAATGTCACACC	1581
OY	1021	TGTGCGCACTGTGTAACCTCAAGAGCAGACAGACAAAGCCCTGTGCACTCCGCAAGCGTG	1080
DB	1582	TGTGCGCACTGTGTAACCTCAAGAGCAGACAGACAAAGCCCTGTGCACTCCGCAAGCGTG	1641
OY	1081	CCCAAGCTGCCCAAGGCTGACAGCCGCTGCTGTTAAGAGACAGCCGCTGTGTACAGGCGTTC	1140
DB	1642	CCCAAGCTGCCCAAGGCTGACAGCCGCTGCTGTTAAGAGACAGCCGCTGTGTACAGGCGTTC	1701

QY	1141	AGTGGGCTGCACCTGTCCGAGAGAGGCTGGAGACAGCAATAGCTCAAGCGTTCCTTC	1200
Db	1702	AGTGGGCTGCACCTGTCCGAGAGAGGCTGGAGACAGCAATAGCTCAAGCGTTCCTTC	1761
QY	1201	TCTCTGATATCAATCAGTTTCATATTACAGCCAGATGGCAGATCCTTACATAGGCTTC	1260
Db	1762	TCTCTGATATCAATCAGTTTCATATTACAGCCAGATGGCAGATCCTTACATAGGCTTC	1821
QY	1261	TCCTCATCAGAGATGCTTTGGAAATCTACAAACCTTCCACTCTGGATGGGACCAAC	1320
Db	1822	TCCTCATCAGAGATGCTTTGGAAATCTACAAACCTTCCACTCTGGATGGGACCAAC	1881
QY	1321	AAGCTATGGCCAGTTCTCCCTCTGTTCAGGAACATAATGGAGCAGACTCCCCGAAACCAAGTCTT	1380
Db	1882	AAGCTATGGCCAGTTCTCCCTCTGTTCAGGAACATAATGGAGCAGACTCCCCGAAACCAAGTCTT	1941
QY	1381	GATATAGGAGGAAGCCAGCATCCCGAAGAGCTGAGACCGGCAGGCTTACAGACAGCAG	1440
Db	1942	GATATAGGAGGAAGCCAGCATCCCGAAGAGCTGAGACCGGCAGGCTTACAGACAGCAG	2001
QY	1441	AGCAAGCGATTTGCATTGCGTTCAGAACCCAGCACGACGAGTGGCACCGCGCCAGAGGTCCCTTTTA	1500
Db	2002	AGCAAGCGATTTGCATTGCGTTCAGAACCCAGCACGACGAGTGGCACCGCGCCAGAGGTCCCTTTTA	2061
QY	1501	TCTTCCACTGCATCGAAGTGGAGGCTGAGAGCAATTACACACACAGCTTCCTTTTGGC	1560
Db	2062	TCTTCCACTGCATCGAAGTGGAGGCTGAGAGCAATTACACACACAGCTTCCTTTTGGC	2121
QY	1561	CTTTTCCACACAGCCAGACACCTACAGAAAGCTGTGCTGGGCTGGGCTTTAAAGGCTGGCAC	1620
Db	2122	CTTTTCCACACAGCCAGACACCTACAGAAAGCTGTGCTGGGCTGGGCTTTAAAGGCTGGCAC	2181
QY	1621	TCGGATATCTTGGGCGCCCCCAGACCTCTACCCCTTCCCTGACACAGCAGCTGGTATTTTGGC	1680
Db	2182	TCGGATATCTTGGGCGCCCCCAGACCTCTACCCCTTCCCTGACACAGCAGCTGGTATTTTGGC	2241
QY	1681	ACAGAGTCTCTACACTTCTACTCTGTGCTCAGCCACTTACCGAGGCACTGTGCACTTACTCT	2301
Db	2302	ACAGAGTCTCTACACTTCTACTCTGTGCTCAGCCACTTACCGAGGCACTGTGCACTTACTCT	2361
QY	1741	GCCTACAGCTGCACACGACCTGCCCCACTTTGGGGGAAACCAAGTCTATTGTGTGGCAGGCGG	1800
Db	2302	GCCTACAGCTGCACACGACCTGCCCCACTTTGGGGGAAACCAAGTCTATTGTGTGGCAGGCGG	2361
QY	1801	CAGAAAGCCAAATGACAGAGCTGACTCCCGGCGGAGCTGGCATGAAAGAGAGAGCCCTTTGAA	1860
Db	2362	CAGAAAGCCAAATGACAGAGCTGACTCCCGGCGGAGCTGGCATGAAAGAGAGAGCCCTTTGAA	2421
QY	1861	AAGCAGTTTAAACGCAGAGAGCTGGCCAATGGAATTTTGGAGAGAGCATCATGTCCAGAGAAC	1920
Db	2422	AAGCAGTTTAAACGCAGAGAGCTGGCCAATGGAATTTTGGAGAGAGCATCATGTCCAGAGAAC	2481
QY	1921	AGGTTCACGGGAAGAGCTGGGGGAAAGTGGGCAAGTCAAGTCTATCTTTTGGGGCAGCATGGAA	1980
Db	2482	AGGTTCACGGGAAGAGCTGGGGGAAAGTGGGCAAGTCAAGTCTATCTTTTGGGGCAGCATGGAA	2541
QY	1981	ATCATTTGAGGTCTCC 1995	
Db	2542	ATCATTTGAGGTCTCC 2556	

RESULT	6
AB052156	
LOCUS	AB052156
DEFINITION	Homo sapiens MKP-7 mRNA for MAPK phosphatase-7, complete cds.
ACCESSION	AB052156
VERSION	AB052156.1 GI:13548676
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1

QY 1501 TCCTCACTGATGAGAGTGGAGAGCTGGAGAGCAATTACACACAGCTTCTTTCCGCG 1560
DB 2064 TCTCCACTGATGAGAGTGGAGAGCTGGAGAGCAATTACACACAGCTTCTTTCCGCG 2123
QY 1561 CTTTCCACGACGACGACGACCTCAGCAAGTCTGTGGCTGGGCTTTAAAGGCTGACAC 1620
DB 2124 CTTTCCACGACGACGACGACCTCAGCAAGTCTGTGGCTGGGCTTTAAAGGCTGACAC 2183
QY 1621 TCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGACACGAGCTGTATTTGGC 1680
DB 2184 TCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGACACGAGCTGTATTTGGC 2243
QY 1681 ACAGAGCTCTCACTTCTACTGCTGCTGACCATCTACGAGGAGGAGCTGCACTTCTCT 1740
DB 2244 ACAGAGCTCTCACTTCTACTGCTGCTGACCATCTACGAGGAGGAGCTGCACTTCTCT 2303
QY 1741 GCCTACAGCTGACGACGACCTGCCCACTTGGCGAGACCAAGTCTATTTCTGCGCAGGCG 1800
DB 2304 GCCTACAGCTGACGACGACCTGCCCACTTGGCGAGACCAAGTCTATTTCTGCGCAGGCG 2363
QY 1801 CAGAGCCAAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATTAAGAGAGCCCTTTGAA 1860
DB 2364 CAGAGCCAAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATTAAGAGAGCCCTTTGAA 2423
QY 1861 AACGAGTTTAAAGCGAAGCTGCCAAATGGAATTTGAGAGAGCATCATGCGAGAAC 1920
DB 2424 AACGAGTTTAAAGCGAAGCTGCCAAATGGAATTTGAGAGAGCATCATGCGAGAAC 2483
QY 1921 AGGTACGAGGAGAGCTGGGGAAGTGGGAGTCACTGCTGCTTTGCGGCGACATGGA 1980
DB 2484 AGGTACGAGGAGAGCTGGGGAAGTGGGAGTCACTGCTGCTTTGCGGCGACATGGA 2543
QY 1981 ATCATTGAGGTCTCC 1995
DB 2544 ATCATTGAGGTCTCC 2558

RESULT 7
AX260340
LOCUS AX260340 3544 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 1 from Patent WO0173059.
ACCESSION AX260340
VERSION AX260340.1 GI:16509303
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Meyers, R.A.
TITLE 3892 and 21117: dual specificity phosphatase molecules and uses
therefor
JOURNAL Patent: WO 0173059-A 1 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Source Location/Qualifiers
1. .3544
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
589..2586
/note="unassigned protein product"
/codon_start=1
/protein_id="CAD10218.1"
/db_xref="GI:16509304"
/db_xref="REMBTREMBL:CAD10218"
/translation="MAHEMTGTOIVTERIALLESGETEKVLIIDSRPVEYNHTLIE
AININCKLMKRLOQDKVLTITELIGHSAHKVDIDCSQVYVYDSSQDVASLSDC
FLTVLGLKESKFSNVLHLAGFAERFRCPLGEGSKSTVPTICPCLPVANIGPT
RLPNTVLGQDVNLKELMQNGIIGVLANASTCPDPDIPESHFLRVVNDSPFCK
ILPMLDSVDFIEKAKASNGCVLHICLAGISRSATLAIAIIMRMMSLDEAYFYKE
KRPTISNPNFGLDLDYEKKIKNTGASPKLKLHLKEKNEPVPVASEGQSGE
TLPSPCADSATSEAGQRPVHPASVPSVQPSLLEDSPLVQALSGHLASDRLED

ORIGIN

Query Match 94.9%; Score 1893; DB 6; Length 3544;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAATCTCAATTTGTAAGAGGTTGGCTGTCTGAA 60
DB 589 ATGGCCCATGAGATGATGGAATCTCAATTTGTAAGAGGTTGGCTGTCTGAA 648
QY 61 AGTGAACGAAAAAGTGTCTGTAATGTAAGCCGCTTTTGTGAATACATATCC 120
DB 649 AGTGAACGAAAAAGTGTCTGTAATGTAAGCCGCTTTTGTGAATACATATCC 708
QY 121 CACATTTGGAAGCCATTAATATCACTGCTCCAGCTTATGAAAGGAGTTGCAACAG 180
DB 709 CACATTTGGAAGCCATTAATATCACTGCTCCAGCTTATGAAAGGAGTTGCAACAG 768
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATGAT 240
DB 769 GACAAAGTGTAAATTAACAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATGAT 828
QY 241 TGCAGTCAAGAGTGTAGTTACGATCAAGAGTCCCAAGTGTGCTCTCTCTTCA 300
DB 829 TGCAGTCAAGAGTGTAGTTACGATCAAGAGTCCCAAGTGTGCTCTCTCTTCA 888
QY 301 GACTGTTTCTACCTGACTCTGCGGTAACTGGAAGAGCTTCACTCTGTCACTG 360
DB 889 GACTGTTTCTACCTGACTCTGCGGTAACTGGAAGAGCTTCACTCTGTCACTG 948
QY 361 CTTCAGAGTGGTTGCTGAGTCTCTGTTGTTTCCCTGAGCTGTGAGAGAAATCC 420
DB 949 CTTCAGAGTGGTTGCTGAGTCTCTGTTGTTTCCCTGAGCTGTGAGAGAAATCC 1008
QY 421 ACTCTAGTCCCTACCTGACATTTCTCAGCTTCTGTTTCCCTGAGCTGTGAGAGAAATCC 480
DB 1009 ACTCTAGTCCCTACCTGACATTTCTCAGCTTCTGTTTCCCTGAGCTGTGAGAGAAATCC 1068
QY 481 CGAATCTTCCCAATCTTATATCTGCTGCGAGCGAGATGCTTCAACAGAGCTGATA 540
DB 1069 CGAATCTTCCCAATCTTATATCTGCTGCGAGCGAGATGCTTCAACAGAGCTGATA 1128
QY 541 CAGCAAAATGGAGATTGTTATGTTAAATGACGATATACCTGTCCAAAGCTGACTTT 600
DB 1129 CAGCAAAATGGAGATTGTTATGTTAAATGACGATATACCTGTCCAAAGCTGACTTT 1188
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 660
DB 1189 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 1248
QY 661 CCGTGTGTGACAAACAGTAGATTTCATTTGAGAAAGCAAAAGCCCTCAATGATGCTT 720
DB 1249 CCGTGTGTGACAAACAGTAGATTTCATTTGAGAAAGCAAAAGCCCTCAATGATGCTT 1308
QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCGGACACATGCTATTCGCTCATCATG 780
DB 1309 CTAGTCACTGTTTAACTGGGATCTCCGCTCGGACACATGCTATTCGCTCATCATG 1368
QY 781 AAGAGATGACATGCTTTTGAAGCTTACAGATTTGTGAGAGAGAGAGAGAGCTTACT 840
DB 1369 AAGAGATGACATGCTTTTGAAGCTTACAGATTTGTGAGAGAGAGAGAGAGCTTACT 1428
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGACATATGAGAGAGAGAGAGAG 900
DB 1429 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGACATATGAGAGAGAGAGAGAG 1488
QY 901 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGACCTGAGAGAGCAAT 960

Db	1489	CAGACTGGAGGCAATCAGAGGCCAAGAGCAAACTCAGACTGCTCACCTGGAGAGCCAAAT	1548
QY	961	GAACTCTGTCCCTGCTGCTTCTCAGAGGCTGGACAGAAAAGCAGAGCGCCCTCAGTCCACC	1020
Db	1549	GAACTCTGTCCCTGCTGCTTCTCAGAGGCTGGACAGAAAAGCAGAGCGCCCTCAGTCCACC	1608
QY	1021	TGAGCGGCACCTGCTGCTGCTCAGAGGAGAGAGCAAAAGGCCCTGTCATCTCCGCCAGCGTG	1080
Db	1609	TGAGCGGCACCTGCTGCTGCTCAGAGGAGAGAGCAAAAGGCCCTGTCATCTCCGCCAGCGTG	1668
QY	1081	CCGACGCTGCCAGCGCTGAGCGCTGCTGCTGTTAGAGAGCAGCCCGCTGTATCAGGCGCTC	1140
Db	1669	CCGACGCTGCCAGCGCTGAGCGCTGCTGCTGTTAGAGAGCAGCCCGCTGTATCAGGCGCTC	1728
QY	1141	AGGCGGCTGCTCACCTGTGCGGCACACAGGCTGAGAAAGCAGCATATAGCTCAAGCGTTCCTT	1200
Db	1729	AGGCGGCTGCTCACCTGTGCGGCACACAGGCTGAGAAAGCAGCATATAGCTCAAGCGTTCCTT	1788
QY	1201	TCTCTGGATATCAAAATCAGATTTCAATTCAGCAGCATGAGCAGATCTCTTACATGGCTTC	1260
Db	1789	TCTCTGGATATCAAAATCAGATTTCAATTCAGCAGCATGAGCAGATCTCTTACATGGCTTC	1848
QY	1261	TCTCTCATCGAAGATGCTTTGGAAATCTTCAAACTTCCACTACTCTTGGATGGAGCCAC	1320
Db	1849	TCTCTCATCGAAGATGCTTTGGAAATCTTCAAACTTCCACTACTCTGATGGAGCCAAC	1908
QY	1321	AAGCTATGCGACGATTTCTCCCTCTTTACAGGAATATGAGAGAGACTCCCGGAACAGTCT	1380
Db	1909	AAGCTATGCGACGATTTCTCCCTCTTTACAGGAATATGAGAGAGACTCCCGGAACAGTCT	1968
QY	1381	GATTAAGAGAGAGCCAGCATCTCCCAAGAGCTGCAGACCGCGCAGGCTTTCAGACAGCAG	1440
Db	1969	GATTAAGAGAGAGCCAGCATCTCCCAAGAGCTGCAGACCGCGCAGGCTTTCAGACAGCAG	2028
QY	1441	AGCAAGCGCATTTGCGTTCAGAACCGACAGCAGTGGCACCGGCCAGAGGTCCTCTTTTA	1500
Db	2029	AGCAAGCGCATTTGCGTTCAGAACCGACAGCAGTGGCACCGGCCAGAGGTCCTCTTTTA	2088
QY	1501	TCTCCACTGTCATCGAAGTGGGAGCGCTGAGAGCAATTACACACACAGCTTCTCTTTTGGC	1560
Db	2089	TCTCCACTGTCATCGAAGTGGGAGCGCTGAGAGCAATTACACACACAGCTTCTCTTTTGGC	2148
QY	1561	CTTTTCCACAGCAGCAGCACACCTCAGCAAAATCTGTGTGGCTTGAAGGCTGGCAC	1620
Db	2149	CTTTTCCACAGCAGCAGCACACCTCAGCAAAATCTGTGTGGCTTGAAGGCTGGCAC	2208
QY	1621	TGCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGCAGTGTATTTTGGC	1680
Db	2209	TGCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGCAGTGTATTTTGGC	2268
QY	1681	ACAGAGTCTCTCAACCTTCTTACTCTGTGCTCAGCCATCTTACGAGGCGAGTGCACGTTACTCT	1740
Db	2269	ACAGAGTCTCTCAACCTTCTTACTCTGTGCTCAGCCATCTTACGAGGCGAGTGCACGTTACTCT	2328
QY	1741	GCTTACAGCTCAGCAGCTGCGCCCATCTTGGCGGAGACCAAGTCAATTTCTGTGGCCAGAGCGG	1800
Db	2329	GCTTACAGCTCAGCAGCTGCGCCCATCTTGGCGGAGACCAAGTCAATTTCTGTGGCCAGAGCGG	2388
QY	1801	CAGAAAGCCAGATGACAGAGCTGACTGCGCGGAGCTGGCATGAGAGAGCCCTTTTGA	1860
Db	2389	CAGAAAGCCAGATGACAGAGCTGACTGCGCGGAGCTGGCATGAGAGAGCCCTTTTGA	2448
QY	1861	AAGCAGTTTAAACCCAGAGAGCTGCCAATATGGAATTTTGAAGAGCATCATGTCAAGAAAC	1920
Db	2449	AAGCAGTTTAAACCCAGAGAGCTGCCAATATGGAATTTTGAAGAGCATCATGTCAAGAAAC	2508
QY	1921	AGGTCACGGGAGAGAGCTGGGGAAAGTGGAGCAGTCACTTCTTTTGGGCGACATGGAA	1980
Db	2509	AGGTCACGGGAGAGAGCTGGGGAAAGTGGAGCAGTCACTTCTTTTGGGCGACATGGAA	2568
QY	1981	ATCATTTAGAGTCTCTC	1995
Db	2569	ATCATTTAGAGTCTCTC	2583

LOCUS	AF506796	3566 bp	mRNA	linear	PRI 30-OCT-2000
DEFINITION	Homo sapiens dual specificity phosphatase 16 (DUSP16) mRNA, complete cds.				
ACCESSION	AF506796				
VERSION	AF506796.1				
KEYWORDS	GI:25573087				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Hoonaeert, I., Marynen, P., Goris, J., Scior, R. and Baens, M.				
TITLE	MAPK phosphatase DUSP16/MKP-7, a candidate tumor suppressor for chromosome region 12p12-13, reduces BCR-ABL-induced transformation Oncogene 22 (49), 7728-7736 (2003)				
JOURNAL	14586399				
PUBMED	2 (bases 1 to 3566)				
REFERENCE	Hoonaeert, I., Marynen, P. and Baens, M.				
AUTHORS	Direct Submission				
TITLE	Submitted (26-APR-2002) Department for Human Genetics-Flinders Interuniversity Institute for Biotechnology (VIB), Katholieke Universiteit Leuven, Herestraat 49, Leuven B-3000, Belgium				
JOURNAL	Location/Qualifiers				
FEATURES	1..3566				
source	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="12"				
	/map="12p13-p12"				
	/tissue_type="bone marrow"				
	1..3566				
gene	/gene="DUSP16"				
	/note="synonym: MKP7"				
	633..2630				
CDS	/gene="DUSP16"				
	/EC_number="3.1.3.48"				
	/EC_number="3.1.3.16"				
	/codon_start=1				
	/product="dual specificity phosphatase 16"				
	/protein_id="AA075120.1"				
	/db_xref="GI:25573088"				
	/translation="MAHEMIGTQIVTERLVALLSGETKVLIDSRPFVEYNTSHILEAINISCLAKMRKLOQDKVLTTELIOHSAKHKVIDCSQKVVVDQSDVASLSDPCFLTVLGLKESFNSVHLAAGFAEFERCFEGLSCGKSTIVPTCSICPELVANIGPGRFLINLVLYGCDRVYINKEIMQONGIGVLNANSTCPMPDPIESHPIRVPNDSFCEKRIIPMLDKSVDFIEKAKNSGCVLHACLAGISRATLIAIYMRKMSLDAVAPVYKRRPLISPNFNLISGLIDYEKKIKNOTGASGSKSLKLHLKPNRPVPAVSEGGQKSTPLSPCADATSEAAAGRPVHPASVPSVPSVOPSLJEDSPVALSGHLHSARLESNKIKRSLDILKSVSYSSASVAAHLHFPSSSEDLLEYKPSITLIDGNKLCQFSPVOELSEDTPEPSDPKEKASIPKKIQTPARPSDSQKRLSHVYRTSSGTACQSLSLPLRRSGSGVEDVHYHFLFGSLSTSOOHLTKSAGLKGWHSDIAPQISTPELTSWYFATSSHFYSAIAIKGSAVSAYVSCSLPTCGDQVYSVRRQKPSDRADSRSMHBSPPFKQKRRRSQGMFEGSINSEKNSRBLCKVSGSSFSGMTIIYVS"				
ORIGIN					
Query Match	94.9%	Score 1893	DB 9	Length 3566	
Best local Similarity	99.9%	Pred. No. 0			
Matches 1993	Conservative 0	Mismatches 2	Indels 0	Gaps 0	
1	ATGGCCCATGATGATGTGAACCTCAATTGTTACTGAGAGGTGGTGGCTCTGCTGGAA	60			
Db	ATGGCCCATGATGATGTGAACCTCAATTGTTACTGAGAGGTGGTGGCTCTGCTGGAA	692			
633	AGTGAACGGAAGAAAGTCTGCTATATGATAGCGGCACATTTGTGAGATACAATATACCC	120			
Qy	AGTGAACGGAAGAAAGTCTGCTATATGATAGCGGCACATTTGTGAGATACAATATACCC	120			
Db	AGTGAACGGAAGAAAGTCTGCTATATGATAGCGGCACATTTGTGAGATACAATATACCC	752			
Qy	CACATTTTGGAGGCATTAATATCACTGCTCCAGCTTATGAAAGCAAGGTTCACAAG	180			

Db 753 CACATTGGAGGCCATTATATCACTGCTCCAGCTTATGAAAGGAAAGTTGCAACAG 812
 Qy 181 GACAAAGTGTATTAACAGAGCTCATCCAGCAATTCAGGAAATATAGGTGACATTAAT 240
 Db 813 GACAAAGTGTATTAACAGAGCTCATCCAGCAATTCAGGAAATATAGGTGACATTAAT 872
 Qy 241 TGCAGTCAGAAAGTTAGTTTACATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 300
 Db 873 TGCAGTCAGAAAGTTAGTTTACATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 932
 Qy 301 GACTGTTTCTCATCTGACTTCTGGGTAACTGAGAAAGCTTCAACTCTGTTCACTG 360
 Db 933 GACTGTTTCTCATCTGACTTCTGGGTAACTGAGAAAGCTTCAACTCTGTTCACTG 992
 Qy 361 CTTCGAGTGGGTTTGTGAGTTCTCTGTTTCTGTTTCTGCTCTGTTGAAAGAAATCC 420
 Db 993 CTTCGAGTGGGTTTGTGAGTTCTCTGTTTCTGTTTCTGCTCTGTTGAAAGAAATCC 1052
 Qy 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCCTTGCTTACCTGTTGCAACATTTGGGCCAAC 480
 Db 1053 ACTCTAGTCCCTACCTGCAATTTCTCAGCCCTTGCTTACCTGTTGCAACATTTGGGCCAAC 1112
 Qy 481 CGAATTCCTCCCAATCTTATCTTGGCTGCGACGAGATGTCCTCAACAGAGCTGATA 540
 Db 1113 CGAATTCCTCCCAATCTTATCTTGGCTGCGACGAGATGTCCTCAACAGAGCTGATA 1172
 Qy 541 CAGCAAAATGGGATTTGTTATGTTTAAATGCGACTATACCTGTCGAAGCTGACTT 600
 Db 1173 CAGCAAAATGGGATTTGTTATGTTTAAATGCGACTATACCTGTCGAAGCTGACTT 1232
 Qy 601 ATCCCGAGTCTCATTTCCGCTGCTGCTGCTGATGAGACAGCTTTGTGAAATTTTG 660
 Db 1233 ATCCCGAGTCTCATTTCCGCTGCTGCTGCTGATGAGACAGCTTTGTGAAATTTTG 1292
 Qy 661 CCGTGGTTGGAACAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCATGATGATGTT 720
 Db 1293 CCGTGGTTGGAACAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCATGATGATGTT 1352
 Qy 721 CTATGTCACCTGTTTACCTGGGATCTCCGCTCCGCCAATCGCTATCGCTTACATGATG 780
 Db 1353 CTATGTCACCTGTTTACCTGGGATCTCCGCTCCGCCAATCGCTTACATGATGATG 1412
 Qy 781 AAGAGATGGAACATGCTTTAGTAGAAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 840
 Db 1413 AAGAGATGGAACATGCTTTAGTAGAAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 1472
 Qy 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGAGTATGAGAAAGATTTAAGAAC 900
 Db 1473 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGAGTATGAGAAAGATTTAAGAAC 1532
 Qy 901 CAGACTGAGACATCAGGGGCCAAAGAGCAAACTCAGCTGCTGCACTTGAAGAACCAAT 960
 Db 1533 CAGACTGAGACATCAGGGGCCAAAGAGCAAACTCAGCTGCTGCACTTGAAGAACCAAT 1592
 Qy 961 GAACTGTCCTGCTGCTCTCAGAGGGTGAAGAAAGAGAGAGCCCTCACTCCACCC 1020
 Db 1593 GAACTGTCCTGCTGCTCTCAGAGGGTGAAGAAAGAGAGAGCCCTCACTCCACCC 1652
 Qy 1021 TGTGCGCAGCTGCTACTCTCAGAGGAGAGAGCAAAAGGCCCTGATCCCGCACGCTG 1080
 Db 1653 TGTGCGCAGCTGCTACTCTCAGAGGAGAGAGCAAAAGGCCCTGATCCCGCACGCTG 1712
 Qy 1081 CCCAGCGTGCAGCGCTGACCGCTGCTGTTAGAGAGACAGCCGCTGTTAAGCGGCTTC 1140
 Db 1713 CCCAGCGTGCAGCGCTGACCGCTGCTGTTAGAGAGACAGCCGCTGTTAAGCGGCTTC 1772
 Qy 1141 AGTGGGCTGCACTGTCGCCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGGCTTCTTC 1200
 Db 1773 AGTGGGCTGCACTGTCGCCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGGCTTCTTC 1832
 Qy 1201 TCTCTGGATATCAAAATCAGTTTCAATTTACAGCCAGATGAGCACTTTCATGAGCTTC 1260
 Db 1833 TCTCTGGATATCAAAATCAGTTTCAATTTACAGCCAGATGAGCACTTTCATGAGCTTC 1892

Qy 1261 TCTCTATCAGAGATGCTTTGGAAATCTACAAACCTTCCACTACTGTGAGTGGACCAAC 1320
 Db 1893 TCTCTATCAGAGATGCTTTGGAAATCTACAAACCTTCCACTACTGTGAGTGGACCAAC 1952
 Qy 1321 AAGCTATGCGAGTTCTCCCTGTTTCCAGAACTATGAGAGCAGACTCCGAAACCACTGCT 1380
 Db 1953 AAGCTATGCGAGTTCTCCCTGTTTCCAGAACTATGAGAGCAGACTCCGAAACCACTGCT 2012
 Qy 1381 GATTAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGACAGCCGAGCTTCAAGACCCAG 1440
 Db 2013 GATTAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGACAGCCGAGCTTCAAGACCCAG 2072
 Qy 1441 AGCAGCGATTCATTCGCTGAGAAACAGCAGAGAGTGGACCGCCAGAGGTCCTTTTA 1500
 Db 2073 AGCAGCGATTCATTCGCTGAGAAACAGCAGAGAGTGGACCGCCAGAGGTCCTTTTA 2132
 Qy 1501 TCTCCACTGCACTGAAGTGGAGCGGTGAGAGCAATTAACAGACAGCTTCTTTTCGGC 1560
 Db 2133 TCTCCACTGCACTGAAGTGGAGCGGTGAGAGCAATTAACAGACAGCTTCTTTTCGGC 2192
 Qy 1561 CTTTCCACAGCCAGCAGACGACTTACAGAAAGTCTGCTGCTGAGCTTAAAGGCTGGCAC 1620
 Db 2193 CTTTCCACAGCCAGCAGACGACTTACAGAAAGTCTGCTGCTGAGCTTAAAGGCTGGCAC 2252
 Qy 1621 TGGGATATCTTGGCCCCCGAGACCTTACCCCTTCCCTGACAGACAGCTGTATTTTGGC 1680
 Db 2253 TGGGATATCTTGGCCCCCGAGACCTTACCCCTTCCCTGACAGACAGCTGTATTTTGGC 2312
 Qy 1681 ACAGAGTCTCACAATCTTACTCTGCTGAGCACTTACAGAGAGAGAGAGAGAGAGAGAG 1740
 Db 2313 ACAGAGTCTCACAATCTTACTCTGCTGAGCACTTACAGAGAGAGAGAGAGAGAGAGAG 2372
 Qy 1741 GCCTACAGCTGACAGCCAGCTGCCCACTTGGAGAGACCAAGTCTATTTGTCGACAGCG 1800
 Db 2373 GCCTACAGCTGACAGCCAGCTGCCCACTTGGAGAGACCAAGTCTATTTGTCGACAGCG 2432
 Qy 1801 CAGAGCCAGTGAACAGACGCTACCTGCGCGGAGAGCTGGCATGAAGAGAGAGAGAGAG 1860
 Db 2433 CAGAGCCAGTGAACAGACGCTACCTGCGCGGAGAGCTGGCATGAAGAGAGAGAGAGAG 2492
 Qy 1861 AAGCAATTTAAAGCAGAGAGTGCCTCAATATGATTTGAGAGAGCATCATGTCAGAGAAC 1920
 Db 2493 AAGCAATTTAAAGCAGAGAGTGCCTCAATATGATTTGAGAGAGCATCATGTCAGAGAAC 2552
 Qy 1921 AGGTCAAGGAGAGAGCTGGGAAAGTGGGAGTCACTTACGTTTTCGGGACAGATGAA 1980
 Db 2553 AGGTCAAGGAGAGAGCTGGGAAAGTGGGAGTCACTTACGTTTTCGGGACAGATGAA 2612
 Qy 1981 ATCATTGAGGCTTCC 1995
 Db 2613 ATCATTGAGGCTTCC 2627

RESULT 9
 AX374994
 LOCUS AX374994 3766 bp DNA linear PAT 01-MAR-2002
 DEFINITION Sequence 17 from Patent WO0210363.
 ACCESSION AX374994
 VERSION AX374994.1 GI:19169826
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS Tang,Y.T., Elliott,V.S., Ramkumar,J., Yao,M.G., Buford,N.,
 Wang,Y.E., Stewart,E.A., Gandhi,A.R., Peterson,C., Lee,E.A.,
 Hatalia,A.J., Lu,D.A., Tribouley,C.M., Griffin,J.A., Baughn,M.R.,
 Yue,H., Warren,B.A., Nguyen,D.B. and Walla,N.K.
 TITLE Protein phosphatases
 Patent: WO 0210363-A 17 07-FEB-2002;
 JOURNAL Incyte Genomics, Inc. (US)

FEATURES		Location/Qualifiers	
source		1..3766	
		/organism="Homo sapiens"	
		/db_type="unassigned DNA"	
		/db_xref="taxon:9606"	
ORIGIN		/note="Incycle ID No: 7480570CB1"	
Query Match 94.9%; Score 1893; DB 6; Length 3766;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	ATGGCCCATGAGATGATTTGAACTCAATTTGTTACTGAGAGTGGTGGCTCTGCTGAA	60
Db	538	ATGGGCCATGAGATGATTTGAACTCAATTTGTTACTGAGAGTGGTGGCTCTGCTGAA	597
Qy	61	AGTGAACGGAAGAAAGTGTCTATTTGATTTAGCCGCGCATTTTGGAAATACATATACATCC	120
Db	598	AGTGAACGGAAGAAAGTGTCTATTTGATTTAGCCGCGCATTTTGGAAATACATATACATCC	657
Qy	121	CACATTTTGAAGCCATTAAATATCAACTGCTCCAAGCTTATGAAAGCGAAGTTGCAACAG	180
Db	658	CACATTTTGAAGCCATTAAATATCAACTGCTCCAAGCTTATGAAAGCGAAGTTGCAACAG	717
Qy	181	GACAAAGTTAATTACAGAGCTATCCAGCATTCAGGAAACATTAAGTTGACATTGAT	240
Db	718	GACAAAGTTAATTACAGAGCTATCCAGCATTCAGGAAACATTAAGTTGACATTGAT	777
Qy	241	TGCAATCAGAAGTTGATTTAGATCAAGCTCCCAAGATGTTGGCTCTCTCTCTCA	300
Db	778	TGCAATCAGAAGTTGATTTAGATCAAGCTCCCAAGATGTTGGCTCTCTCTCTCA	837
Qy	301	GACTGTTTTCACCTGATCTTCTGGGTAACTGAGAAAGACTTCAACTCTGTTCACTG	360
Db	838	GACTGTTTTCACCTGATCTTCTGGGTAACTGAGAAAGACTTCAACTCTGTTCACTG	897
Qy	361	CTTCAGAGTGGGTTTGCAGAGTCTCTCGTTGTTTCCCTGGCTCTGTGAGAAATCC	420
Db	898	CTTCAGAGTGGGTTTGCAGAGTCTCTCGTTGTTTCCCTGGCTCTGTGAGAAATCC	957
Qy	421	ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTACTCTGTTGCAACATTTGGGCCAAC	480
Db	958	ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTACTCTGTTGCAACATTTGGGCCAAC	1017
Qy	481	CGAATTTTCCCAATCTTTATCTTTGGCTGCGCAGCGAGATGCTCTCAACAGAGCTGAT	540
Db	1018	CGAATTTTCCCAATCTTTATCTTTGGCTGCGCAGCGAGATGCTCTCAACAGAGCTGAT	1077
Qy	541	CAGCAGATGGGATTTGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT	600
Db	1078	CAGCAGATGGGATTTGTTATGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT	1137
Qy	601	ATCCCCGAGTCTCATATTTCTGCGGTGCTGTGATGACAGCTTTTGTGAGAAATTTTG	660
Db	1138	ATCCCCGAGTCTCATATTTCTGCGGTGCTGTGATGACAGCTTTTGTGAGAAATTTTG	1197
Qy	661	CCGTGGTTGGACAAATCAGTAGTTTCATTTGAGAAAGCAAAGCCTCCATGATGTGT	720
Db	1198	CCGTGGTTGGACAAATCAGTAGTTTCATTTGAGAAAGCAAAGCCTCCATGATGTGT	1257
Qy	721	CTAGTCACTGTTTACCTGGGATCTCCGCTCGGCCACCATGCTATGCGCTCAATCATG	780
Db	1258	CTAGTCACTGTTTACCTGGGATCTCCGCTCGGCCACCATGCTATGCGCTCAATCATG	1317
Qy	781	AAGAGATGAGCATGCTTTTAAATGAGCTTACAGATTTTGAAGAAAGAAAGACCTACT	840
Db	1318	AAGAGATGAGCATGCTTTTAAATGAGCTTACAGATTTTGAAGAAAGAAAGACCTACT	1377
Qy	841	ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCGAGCTAATGAGAAAGATTTAGAC	900
Db	1378	ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCGAGCTAATGAGAAAGATTTAGAC	1437
Qy	901	CAGACTGAGCATCAAGGCCAAAGCAAACCTCAAGCTGTGCACTTGAAGACCAAT	960

Db	1438	CAGACTGAGCATCAAGGCCAAAGCAAACCTCAAGCTGTGCACTTGAAGACCAAT	1497
Qy	961	GAACCTGTCCCTGTCTCTCAGAGGTGACAGAAAGAGGACAGCCCTCAGTCCACCC	1020
Db	1498	GAACCTGTCCCTGTCTCTCAGAGGTGACAGAAAGAGGACAGCCCTCAGTCCACCC	1557
Qy	1021	TGTGCCGACTTGTCTACTCTCAGAGGACAGGACAAAGGCCGTGATCCCGCAGCGTG	1080
Db	1558	TGTGCCGACTTGTCTACTCTCAGAGGACAGGACAAAGGCCGTGATCCCGCAGCGTG	1617
Qy	1081	CCGAGGTCGCCAGGCTGACAGCCGTCCGTGTTAGAGACAGCCCGCTGTACAGGCGTC	1140
Db	1618	CCGAGGTCGCCAGGCTGACAGCCGTCCGTGTTAGAGACAGCCCGCTGTACAGGCGTC	1677
Qy	1141	AGTGGCTGCACTGTCCGACAGAGGCTGGAAGACACAAATTAAGCTCAAGGTTCCCTC	1200
Db	1678	AGTGGCTGCACTGTCCGACAGAGGCTGGAAGACACAAATTAAGCTCAAGGTTCCCTC	1737
Qy	1201	TCTCTGATATCAATCAAGTTTCAATTTCAAGCAGATGGGACATCTTACATGCTTC	1260
Db	1738	TCTCTGATATCAATCAAGTTTCAATTTCAAGCAGATGGGACATCTTACATGCTTC	1797
Qy	1261	TCCCATCAGAAGATGCTTTGGAATACAAACCTTCCACTTACTTGTGATGGGACCAAC	1320
Db	1798	TCCCATCAGAAGATGCTTTGGAATACAAACCTTCCACTTACTTGTGATGGGACCAAC	1857
Qy	1321	AAGCTATCCAGTTCTCCCTGTTCAGGAACTATCCAGAGCAGATCCCGAAACAGTCT	1380
Db	1858	AAGCTATCCAGTTCTCCCTGTTCAGGAACTATCCAGAGCAGATCCCGAAACAGTCT	1917
Qy	1381	GATTAAGAGGAAGCCAGATCCCAAGAGCTGAGACCCGACAGGCTTACAGACGACAG	1440
Db	1918	GATTAAGAGGAAGCCAGATCCCAAGAGCTGAGACCCGACAGGCTTACAGACGACAG	1977
Qy	1441	AGCAAGCATTTGCAATTCGATCAGAACAGCAGAGTGGCAACCGCCAGAGGTCCTTTA	1500
Db	1978	AGCAAGCATTTGCAATTCGATCAGAACAGCAGAGTGGCAACCGCCAGAGGTCCTTTA	2037
Qy	1501	TCTCAGCTGATCGAATGGGAGGCTGAGAGCAATTAACACACAGACTTCTTTTCGGC	1560
Db	2038	TCTCAGCTGATCGAATGGGAGGCTGAGAGCAATTAACACACAGACTTCTTTTCGGC	2097
Qy	1561	CTTTCACACAGCCAGCAGCCTCAGAGATGCTGGGCTTGAAGGCTGGCAC	1620
Db	2098	CTTTCACACAGCCAGCAGCCTCAGAGATGCTGGGCTTGAAGGCTGGCAC	2157
Qy	1621	TGGATATCTTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGC	1680
Db	2158	TGGATATCTTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGC	2217
Qy	1681	ACAGATGCTCAACACTTTTACTCTGCTCAGGCCATCTAAGAGGACGTGCAGTTACTCT	1740
Db	2218	ACAGATGCTCAACACTTTTACTCTGCTCAGGCCATCTAAGAGGACGTGCAGTTACTCT	2277
Qy	1741	GCTTACAGCTGACAGCAGCTGCCCACTTGGGAGACCAAGTCTATCTGTGGGCGGGCGG	1800
Db	2278	GCTTACAGCTGACAGCAGCTGCCCACTTGGGAGACCAAGTCTATCTGTGGGCGGGCGG	2337
Qy	1801	CAGAGCGAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATGAAGAGACCCCTTTGAA	1860
Db	2338	CAGAGCGAAGTACAGAGCTGACTGCGCGGAGAGCTGGCATGAAGAGACCCCTTTGAA	2397
Qy	1861	AAGCAGTTTAAACGAGAGCTGCCAAATGGAATTTGAGAGACATCTATCAAGAAC	1920
Db	2398	AAGCAGTTTAAACGAGAGCTGCCAAATGGAATTTGAGAGACATCTATCAAGAAC	2457
Qy	1921	AGGTCAACGGAAGAGCTGGGGAAGAGTGGCAGTCACTTGTGCGGCGAGATGGA	1980
Db	2458	AGGTCAACGGAAGAGCTGGGGAAGAGTGGCAGTCACTTGTGCGGCGAGATGGA	2517
Qy	1981	ATCATTGAGGTCTCC	1995

Db 2518 ATCATTGAGGCTCC 2532

RESULT 10

BD171157 4790 bp DNA linear PAT 17-JAN-2003

LOCUS Novel gene and protein encoded thereby.

DEFINITION BD171157

ACCESSION BD171157

VERSION BD171157.1 GI:27876969

KEYWORDS WO 02052005-A/13.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 4790)

AUTHORS Ohara,O., Nagaue,T. and Nakajima,D.

TITLE Novel gene and protein encoded thereby

JOURNAL Patent: WO 02052005-A 13 04-JUL-2002;

KAZUSA DNA RESEARCH INSTITUTE,OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA

COMMENT OS Homo sapiens (human)

PN WO 02052005-A/13

PD 04-JUL-2002

PR 20-DEC-2001 WO 2001JP011217

PI 22-DEC-2000 JP 00P 389742

PC OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA

CC C12N15/12,C07K14/47

NC Novel gene and protein encoded thereby

FT Key Location/Qualifiers

Location/Qualifiers (184).. (2178).

FEATURES

source 1..4790

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 94.9%; Score 1893; DB 6; Length 4790;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGATGATTTGAACCTCAATTGTTACTGAGAGTGTGGCTGCTGCTGGA 60

DB 184 ATGGCCCATGATGATTTGAACCTCAATTGTTACTGAGAGTGTGGCTGCTGCTGGA 243

QY 61 AGTGGAAACGAAAGTGTCTAATGATAGCCGACATTTGGAATACATATACATCC 120

DB 244 AGTGGAAACGAAAGTGTCTAATGATAGCCGACATTTGGAATACATATACATCC 303

QY 121 CACATTTTGGAGCATTAATATCACTGCTCCAGCTTATGAAGCGAAGTTGGCAACAG 180

DB 304 CACATTTTGGAGCATTAATATCACTGCTCCAGCTTATGAAGCGAAGTTGGCAACAG 363

QY 181 GACAAAGTGAATTAACAGAGCTATCCAGATTCAGCGAAACATAAGTTGACATTTGAT 240

DB 364 GACAAAGTGAATTAACAGAGCTATCCAGATTCAGCGAAACATAAGTTGACATTTGAT 423

QY 241 TGCAGTCAGAGGTTGATGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300

DB 424 TGCAGTCAGAGGTTGATGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 483

QY 301 GACTGTTTCTCACTGATCTTCTGGGTAAACCTGGAGAAAGCTTCAACTCTGTTACCTG 360

DB 484 GACTGTTTCTCACTGATCTTCTGGGTAAACCTGGAGAAAGCTTCAACTCTGTTACCTG 543

QY 361 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTGTTCTGCTGAGCCCTCTGGAAGGAAATCC 420

DB 544 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTGTTCTGCTGAGCCCTCTGGAAGGAAATCC 603

QY 421 ACTCTAGTCCCTACCTGATCTTCTCAAGCTTGTACCTGTTGCAACATTTGGCCCAAC 480

DB 604 ACTCTAGTCCCTACCTGATCTTCTCAAGCTTGTACCTGTTGCAACATTTGGCCCAAC 663

QY 481 CGAATTCCTCCCAATCTTATATCTTGGCTGCAGAGATGCTCTCAACAGAGCTGATA 540

DB 664 CGAATTCCTCCCAATCTTATATCTTGGCTGCAGAGATGCTCTCTCAACAGAGCTGATA 723

QY 541 CAGCAGAAATGGATGATGATATGTTTAAATGATCCAGCTATACCTGTCCAAAGCTGACTTT 600

DB 724 CAGCAGAAATGGATGATGATATGTTTAAATGATCCAGCAATACCTGTCCAAAGCTGACTTT 783

QY 601 ATCCCGAGTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

DB 784 ATCCCGAGTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843

QY 661 CCGTGGTGGACAAATCAGATGATTTCAATGAGAAAGCAAAAGCTCCCAATGATGATGTT 720

DB 844 CCGTGGTGGACAAATCAGATGATTTCAATGAGAAAGCAAAAGCTCCCAATGATGATGTT 903

QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCAACATGCTATGCTTACATCATG 780

DB 904 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCAACATGCTATGCTTACATCATG 963

QY 781 AAGAGATGACATGTTCTTAAATGATGAACTTAAAGATTTGTGAAAGAAAGAAAGCTTACT 840

DB 964 AAGAGATGACATGTTCTTAAATGATGAACTTAAAGATTTGTGAAAGAAAGAAAGCTTACT 1023

QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGGAATGAGAAAGATTTAAGAAC 900

DB 1024 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGGAATGAGAAAGATTTAAGAAC 1083

QY 901 CAGACTGAGCATCAGGCGCAAGAAAGCAAACTCACTGCTGCACTGAGAAAGCAAAAT 960

DB 1084 CAGACTGAGCATCAGGCGCAAGAAAGCAAACTCACTGCTGCACTGAGAAAGCAAAAT 1143

QY 961 GAACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

DB 1144 GAACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1203

QY 1021 TGTGCCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

DB 1204 TGTGCCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1263

QY 1081 CCCAGCGTCCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGC 1140

DB 1264 CCCAGCGTCCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGC 1323

QY 1141 AGTGGCTGCACCTGTCGAGACAGAGCTGGAAGACAGCAATTAAGCTCAAGCTTCTCTTC 1200

DB 1324 AGTGGCTGCACCTGTCGAGACAGAGCTGGAAGACAGCAATTAAGCTCAAGCTTCTCTTC 1283

QY 1201 TCTCTGATATCAATCAAGTTTCAATTAATGAGCAGATGAGCAGATGAGCAGATGAGCAG 1260

DB 1384 TCTCTGATATCAATCAAGTTTCAATTAATGAGCAGATGAGCAGATGAGCAGATGAGCAG 1443

QY 1261 TCTCTATCAGAAAGATCTTGTGAATCTCAAACTTCTCACTCTGATGAGGACCAAC 1320

DB 1444 TCTCTATCAGAAAGATCTTGTGAATCTCAAACTTCTCACTCTGATGAGGACCAAC 1503

QY 1321 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATGAGAGCACTCCGAAACCAAGTCTT 1380

DB 1504 AAGCTATGCCAGTTCTCCCTGTTCAAGAACTATGAGAGCACTCCGAAACCAAGTCTT 1563

QY 1381 GATTAAGAGAAAGCCAGCAATCCCAAGAAAGCTGCAACCGCGGCTTCAAGCAGCAG 1440

DB 1564 GATTAAGAGAAAGCCAGCAATCCCAAGAAAGCTGCAACCGCGGCTTCAAGCAGCAG 1623

QY 1441 AGCAAGCAATGATTCGCTGCAAGAACAGCAGAGTGGCAACCGCGGCTTCAAGCAGCAG 1500

DB 1624 AGCAAGCAATGATTCGCTGCAAGAACAGCAGAGTGGCAACCGCGGCTTCAAGCAGCAG 1683

QY 1501 TCTCACTGATGAGAGTGGAGAGCTGAGAGCAATTAACACAGCTTCTTTTTCGCG 1560

DB 1684 TCTCACTGATGAGAGTGGAGAGCTGAGAGCAATTAACACAGCTTCTTTTTCGCG 1743

QY 1561 CTTTCAACAGCCAGAGACACCTCAGAAAGTCTGAGCTGAGGCTTAAAGGCTGAGCAG 1620

Db	1744	CTTTCACACAGCAGACAGCACTTACGAAAGTCGTGGCTTGGGCTTAAAGGCTGGAC	1803
Qy	1621	TCGGATATCTTGGAGCCCCAGACCTCTACCCCTTCCCTGACACAGAGCGGATTTTCC	1680
Db	1804	TGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGAGCTGGATTTTCC	1863
Qy	1681	ACAGAGTCCTCACTTCTACTCTTCCCTCAGCCATCTACGAGGACGAGTCCAGTTACTCT	1740
Db	1864	ACAGAGTCCTCACTTCTACTCTTCCCTCAGCCATCTACGAGGACGAGTCCAGTTACTCT	1923
Qy	1741	GCTTACAGCTGACGCGACGTGCTCCACTTGGCGAGACCAAGTCTATTCTGTGGCAGCGG	1800
Db	1924	GCTTACAGCTGACGCGACGTGCTCCACTTGGCGAGACCAAGTCTATTCTGTGGCAGCGG	1983
Qy	1801	CAGAAAGCAAGTGAAGAGCTGACCTGGGGGGAGCTGGCATGAAAGAGAGCCCTTTGAA	1866
Db	1984	CAGAAAGCAAGTGAAGAGCTGACCTGGGGGGAGCTGGCATGAAAGAGAGCCCTTTGAA	2043
Qy	1861	AAGCAGTTTAAACGCAGAAAGCTGCCAAATGGAATTTGGAGAGCATCATGTCCAGAGAC	1920
Db	2044	AAGCAGTTTAAACGCAGAAAGCTGCCAAATGGAATTTGGAGAGCATCATGTCCAGAGAC	2103
Qy	1921	AGGTCCACGGGAAGAGCTGGGGGAAAGTGGGACATGATCTACTTTTGGGCGAGCATGGAA	1980
Db	2104	AGGTCCACGGGAAGAGCTGGGGGAAAGTGGGACATGATCTACTTTTGGGCGAGCATGGAA	2163
Qy	1981	ATCATTTAGAGTCTCC	1995
Db	2164	ATCATTTAGAGTCTCC	2178

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	JOURNAL	COMMENT
BD183422	BD183422	Novel genes and proteins encoded by the genes.	BD183422	BD183422.1	GI:31875622	JP 2002345492-A/135.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	Chara,O., Nagase,T. and Nakajima,D. Novel genes and proteins encoded by the genes Patent: JP 2002345492-A 135 03-DEC-2002; KAZUO DNA RESEARCH INSTITUTE	OS Homo sapiens (human) PN JP 2002345492-A/135 PD 03-DEC-2002 PF 26-FEB-2002 JP 2002049009 PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA PC C12N15/09,C07K14/47//A61K31/711,A61K38/00,A61P25/00, PC A61P25/14, PC A61P25/18,A61P35/00,C12N15/00,A61K37/02 CC Novel genes and proteins encoded by the genes FH key location/Qualifiers FT CDS (184). (2178).	

```

FEATURES      . . . . . 1.84 / . . (2.1.0) .
SOURCE
1. .4790      Location/Qualifiers
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"

```

ORIGIN

Query Match	94.9%	Score 1893;	DB 6;	Length 4790;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1993; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY 1 ATGGCCCATGAGATGATTGGAATCAATTGTTACTAGAGAGTTGTGGCTCTGTCGAA 60
bb 184 ATGGCCCATGAGATGATTGGAATCAATTGTTACTAGAGAGTTGTGGCTCTGTCGAA 243

QY	61	AGTGAACGGAAAAAGTGTCTGTAATTAATGATACCGGCACTTTGTGGAATACATATACATCC	120
Db	244	AATGGAAACGGAAAAAGTGTCTGTAATTAATGATACCGGCACTTTGTGGAATACATATACATCC	303
QY	121	CACATTTTGGAAAGCCATTATATATCAACTGTCTCAAGCTTATAGAAGGAAGTTGCACACG	180
Db	304	CACATTTTGGAAAGCCATTATATATCAACTGTCTCAAGCTTATAGAAGGAAGTTGCACACG	363
QY	181	GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAA CATTAAGTTGACATTTGAT	240
Db	364	GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAA CATTAAGTTGACATTTGAT	423
QY	241	TGCAGTCGAAAGTTGTGATTATACATCAAAAGCTCCCAAGATGTTGCCCTCTCTCTTCA	300
Db	424	TGCAGTCGAAAGTTGTGATTATACATCAAAAGCTCCCAAGATGTTGCCCTCTCTCTTCA	483
QY	301	GACTGTTTCTCACTGTAACCTTGTGGGTAACTCGAAGAGAGCTTCAACTCTGTTCACCTG	360
Db	484	GACTGTTTCTCACTGTAACCTTGTGGGTAACTCGAAGAGAGCTTCAACTCTGTTCACCTG	543
QY	361	CTTGCAGGTGGGTTTGTCTGAGTTCTCTCGTTGTTTTCCCTGGCTCTGTGAAAGAAAATCC	420
Db	544	CTTGCAGGTGGGTTTGTCTGAGTTCTCTCGTTGTTTTCCCTGGCTCTGTGAAAGAAAATCC	603
QY	421	ACTCTAATGCCCACTGCACTTCTCAAGCTTGTCTTAACTCTGTGTCCAACTTTGGGCCAAC	480
Db	604	ACTCTAATGCCCACTGCACTTCTCAAGCTTGTCTTAACTCTGTGTCCAACTTTGGGCCAAC	663
QY	481	CGAATTTCTCCCAATCTTTATCTTGGCTGCAGACAGATATGCTCTCAACAGAGAGCTGATA	540
Db	664	CGAATTTCTCCCAATCTTTATCTTGGCTGCAGACAGATATGCTCTCAACAGAGAGCTGATG	723
QY	541	CAGCAAAATGGAGTTGTTATGTGTAAATGCGACGTAATCCTGTCCAAAGCTGACTTT	600
Db	724	CAGCAAAATGGAGTTGTTATGTGTAAATGCGACCAATTAATCCTGTCCAAAGCTGACTTT	783
QY	601	ATCCCCGAGTCTCATTTTCTCGCTGTGCTGTGGAATGACAGTTTGTGAGAAATTTTGG	660
Db	784	ATCCCCGAGTCTCATTTTCTCGCTGTGCTGTGGAATGACAGTTTGTGAGAAATTTTGG	843
QY	661	CCGTGTGTGGAACAAATCAGTAGATTTCATTGAGAAAGCAAAAAGCTCCAAATGATGTGTT	720
Db	844	CCGTGTGTGGAACAAATCAGTAGATTTCATTGAGAAAGCAAAAAGCTCCAAATGATGTGTT	903
QY	721	CTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGCACTATCGCTATCGCTCAATCAATG	780
Db	904	CTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGCACTATCGCTATCGCTCAATCAATG	963
QY	781	AAGAGGATGGAACATGCTCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAGACCTACT	840
Db	964	AAGAGGATGGAACATGCTCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAGACCTACT	1023
QY	841	ATATCTCCAAACTTCAATTTTCTGGGCAACTCTCTGCACTATGAGAAAGATTTAAGAAC	900
Db	1024	ATATCTCCAAACTTCAATTTTCTGGGCAACTCTCTGCACTATGAGAAAGATTTAAGAAC	1083
QY	901	CAGACTGAGCATCAAGGCAAAAGCAAACTCAAGCTGTGCACCTGTGAGAAAGCCAAAT	960
Db	1084	CAGACTGAGCATCAAGGCAAAAGCAAACTCAAGCTGTGCACCTGTGAGAAAGCCAAAT	1143
QY	961	GAACTGTGCCCTGTGCTCTCAAGAGGTGGAACAGAAAAGCGAGCGCCCTCAATGTCAACC	1022
Db	1144	GAACTGTGCCCTGTGCTCTCAAGAGGTGGAACAGAAAAGCGAGCGCCCTCAATGTCAACC	1203
QY	1021	TGTGCGCACTGTGCTACTCAGAGGACAGACAAAGGCCGTGTGATCCCGCAGACGTG	1080
Db	1204	TGTGCGCACTGTGCTACTCAGAGGACAGACAAAGGCCGTGTGATCCCGCAGACGTG	1263
QY	1081	CCAGCGGTGCCAGCGTGACAGCGGTGTGTGTAAGGACAGCGCCGTGTGATCAGCGCTC	1140
Db	1264	CCAGCGGTGCCAGCGTGACAGCGGTGTGTGTAAGGACAGCGCGGTGTGATCAGCGCTC	1323

QY 1141 AGGAGGCTGACCTGTCCGAGACAGGCTGGAAGACGAAATAGCTCAAGCTTCCTTC 1200
 DB 1324 AGTGGGCTGCACCTGTCCGAGACAGGCTGGAAGACGAAATAGCTCAAGCTTCCTTC 1383
 QY 1201 TCTCTGGATATCAAAATCAGTTTCAATATTCAGCCAGACATGCGATCTTACATGGCTTC 1260
 DB 1384 TCTCTGGATATCAAAATCAGTTTCAATATTCAGCCAGACATGCGATCTTACATGGCTTC 1443
 QY 1261 TCTCTCATCAGAAATGCTTTGGAAATCTACAAACCTTCCACTCTCTGGATGGAGCCAC 1320
 DB 1444 TCTCTCATCAGAAATGCTTTGGAAATCTACAAACCTTCCACTCTCTGGATGGAGCCAC 1503
 QY 1321 AAGCTATGCCAGATTCTCCCTGTTACAGAACTATCGAGAGAGCTCCCGAAACAGTCTT 1380
 DB 1504 AAGCTATGCCAGATTCTCCCTGTTACAGAACTATCGAGAGAGCTCCCGAAACAGTCTT 1563
 QY 1381 GATTAAGAGAGAGACGACATCTCCCAAGAACTGACAGCCCGACGCTTTCAGACAGCAG 1440
 DB 1564 GATTAAGAGAGAGACGACATCTCCCAAGAACTGACAGCCCGACGCTTTCAGACAGCAG 1623
 QY 1441 AGCAAGGATTCGATTCGGTCAAGACGACAGAGCTGCGACCGCCAGAGCTCCCTTTTA 1500
 DB 1624 AGCAAGGATTCGATTCGGTCAAGACGACAGAGCTGCGACCGCCAGAGCTCCCTTTTA 1583
 QY 1501 TCTCCACTGCATGGAAGTGGAGCGTGGAGAGCAATTACACACGAGCTTCCTTTCCGC 1560
 DB 1684 TCTCCACTGCATGGAAGTGGAGCGTGGAGAGCAATTACACACGAGCTTCCTTTCCGC 1743
 QY 1561 CTTTCCACGACGACGACGACCTCAGCAAGTCTGCTGGGCTTGAAGGCTGGAC 1620
 DB 1744 CTTTCCACGACGACGACGACCTCAGCAAGTCTGCTGGGCTTGAAGGCTGGAC 1803
 QY 1621 TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACGACGCTGTAATTTGGC 1680
 DB 1804 TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACGACGCTGTAATTTGGC 1863
 QY 1681 ACAGAGCTCTCAGACTCTACTGCTCTGACCATCAGAGGAGGAGTGCAGTACTCTC 1740
 DB 1864 ACAGAGCTCTCAGACTCTACTGCTCTGACCATCAGAGGAGGAGTGCAGTACTCTC 1923
 QY 1741 GCCTACAGCTGACGACGACCTGACCTTGGCGAGACCAATCTATTCTGTGCGCAGGCG 1800
 DB 1924 GCCTACAGCTGACGACGACCTGACCTTGGCGAGACCAATCTATTCTGTGCGCAGGCG 1983
 QY 1801 CAGAAAGCAAGTGAAGAGCTGACTGCGCGGAGAGCTGCGATGAAGAGAGCCCTTTGAA 1860
 DB 1984 CAGAAAGCAAGTGAAGAGCTGACTGCGCGGAGAGCTGCGATGAAGAGAGCCCTTTGAA 2043
 QY 1861 AAGCAGTTTAAAGCGCAAGCTGCCAAATGGAAATTTGGAAGAGCATCATGTCAAGAAC 1920
 DB 2044 AAGCAGTTTAAAGCGCAAGCTGCCAAATGGAAATTTGGAAGAGCATCATGTCAAGAAC 2103
 QY 1921 AGGTCACGGAAGAGCTGGGGAAGAGTGGGAGTCACTGCTGCTTTGCGGCGACATGGAA 1980
 DB 2104 AGGTCACGGAAGAGCTGGGGAAGAGTGGGAGTCACTGCTGCTTTGCGGCGACATGGAA 2163
 QY 1981 ATCATTGAGGTCTCC 1995
 DB 2164 ATCATTGAGGTCTCC 2178

RESULT 12
 AB051487 4790 bp mRNA linear PRI 07-FEB-2001
 LOCUS AB051487
 DEFINITION Homo sapiens mRNA for KIAA1700 protein, partial cds.
 ACCESSION AB051487
 VERSION AB051487.1 GI:12697944
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases)

AUTHORS Nagase, T., Kikuno, R., Hattori, A., Kondo, Y., Okumura, K. and Ohara, O.
 TITLE Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
 JOURNAL DNA Res. 7 (6), 347-355 (2000)
 MEDLINE 21082932
 PUBMED 11214970
 REFERENCE 2 (bases 1 to 4790)
 AUTHORS Ohara, O., Nagase, T. and Kikuno, R.
 TITLE Direct Submission
 JOURNAL Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: cdainfo@kazusa.or.jp, URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913, Fax: 81-438-52-3914)

FEATURES
 source
 1..4790
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="FJ15353"
 /note="vector: pBluescriptII SK plus"
 1..4790
 /gene="KIAA1700"
 <109..2181
 /gene="KIAA1700"
 /note="Start codon is not identified."
 /product="KIAA1700 protein"
 /protein_id="BAB21791.1"
 /db_xref="GI:12697945"
 /translation="A P E H T T T S S M O I K E G G K R G L I Y V M A H E M I C T O I V T R I V A L L E S G E K V L I D S R P E V I N T S I L E H I N I N G S K M R R I L Q D K V L T E L I Q S A K R V D I D C S Q V V Y D Q S S O D V A S L S D C F L V L L L K L E K S F N S V H L A G F A E P S R C P F G L C E G K C T L V P T C I S O P C L P V A N I G P R I L P N L Y L G C O R D V L N E L M Q O N G I G V L N A S N T C P K P D F I P E S H F L R V P V D S F C E K L I P W L D S V D F I E K A K A S N G C V L V H C L A G I S R A T I A I Y I M K R M D S L D E A V R F V K E R P T I S N F N L A G O L L Y E K K I K N O T G A S G S K I K L I H E K P N E P P V A S R A D S G A C K S T P I S P R A D A T S A A G O R P V H P A S V S P E V P L E D S P L V O L S G I H S A D R L E D S N K L R E P S L D I K S V S I S A S M A L H F S S E D A L E Y K P S T T L L G T N K L C O P S P V Q E L S E Q T P E P T S P U K E B A S I P K L I Q T A R P S O S K R L H S V T S S E G T A Q R S L S P L H R S G V E D N V H T S F L G S T S Q O H L T S A G L G I K G W H S D I L A P O T S P S L T S W Y F A T E S H F Y S A A I Y G S A V S A Y S C S Q L P T C G D V Y S V R R K P S D R A D S R S W H E S P E K Q P R R S C Q M E F G S I M S E N R S R E L K V G S S S F S G S M E I L I E V S"

ORIGIN
 Query Match 94.9%; Score 1893; DB 9; Length 4790;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGTGAATTGGAATCTCAATTTGTTACTGAGAGTTGGTGGCTCTGCTGAA 60
 DB 184 ATGGCCCATGAGTGAATTGGAATCTCAATTTGTTACTGAGAGTTGGTGGCTCTGCTGAA 243
 QY 61 AGTGAACGGAAAAAAGTGTGCTAAATGTAAGCCGGCCATTGTTGGAATACATATACATCC 120
 DB 244 AGTGAACGGAAAAAAGTGTGCTAAATGTAAGCCGGCCATTGTTGGAATACATATACATCC 303
 QY 121 CACATTTTGAAGCCATATATCAACTGCTCCAACTTATGAAAGGAGTTGCAACAG 180
 DB 364 GACAAAGTGTATATCAAGCTCATCCAGCTTACGCAAAACATTAAGTTGACATTGAT 423
 QY 241 TGAGTCAGAGTTGTATTTACATCACTCAAGCTCCCAAGATGGTCCCTCTCTTCA 300
 DB 424 TGAGTCAGAGTTGTATTTACATCACTCAAGCTCCCAAGATGGTCCCTCTCTTCA 483
 QY 301 GACTGTTTCTCACTGTACTTCTGGGTAACTGAGAAAGCTTCAACTCTGTTCACTG 360
 DB 484 GACTGTTTCTCACTGTACTTCTGGGTAACTGAGAAAGCTTCAACTCTGTTCACTG 543

QY 361 CTTGAGGTGGGTTTGTGAGTCTCTGCTGTTTCCCTGGCTCTGTGAAGAAATCC 420
|
|
|
Db 544 CTTGAGGTGGGTTTGTGAGTCTCTGCTGTTTCCCTGGCTCTGTGAAGAAATCC 603
|
|
|
QY 421 ACTTAGTCCCTACCTGATTTCTCAGCCTTGCTTACCTGTGGCAATTCATTTGGGCAAC 480
|
|
|
Db 604 ACTTAGTCCCTACCTGATTTCTCAGCCTTGCTTACCTGTGGCAATTCATTTGGGCAAC 663
|
|
|
QY 481 GGAATTTCTCCCAATCTTATCTTGGCTGCGCAGAGATGTCTTCAACAAGAGCTGATA 540
|
|
|
Db 664 GGAATTTCTCCCAATCTTATCTTGGCTGCGCAGAGATGTCTTCAACAAGAGCTGATA 723
|
|
|
QY 541 CAGCAGATGGATGTTATGTATGTAAATGCGAGCTTACCTGTGCAAAAGCTGACCTT 600
|
|
|
Db 724 CAGCAGATGGATGTTATGTATGTAAATGCGAGCAATACCTGTGCAAAAGCTGACCTT 783
|
|
|
QY 601 ATCCCGAGTCTCATTTCTCGGCTGCTGTGAATGA CAGCTTTTGTGAAATTTTG 660
|
|
|
Db 784 ATCCCGAGTCTCATTTCTCGGCTGCTGTGAATGA CAGCTTTTGTGAAATTTTG 843
|
|
|
QY 661 CGGTGGTGGACAATCAGTAGATTCTTATGGAAGAAAGCAAAAGCTTCAATGATGTGT 720
|
|
|
Db 844 CGGTGGTGGACAATCAGTAGATTCTTATGGAAGAAAGCAAAAGCTTCAATGATGTGT 903
|
|
|
QY 721 CTAGTGACTGTGTTAGCTGGGATCTCCGCTCCGCAACATCGCTATCGCTACATCATG 780
|
|
|
Db 904 CTAGTGACTGTGTTAGCTGGGATCTCCGCTCCGCAACATCGCTATCGCTACATCATG 963
|
|
|
QY 781 AAGAGATGAGCACTGCTTTTGAATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
|
|
|
Db 964 AAGAGATGAGCACTGCTTTTGAATGAAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1023
|
|
|
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTCTGACTATGAGAAAGAAATTAAGAC 900
|
|
|
Db 1024 ATATCTCCAACTTCAATTTTCTGGGCAACTCTCTGACTATGAGAAAGAAATTAAGAC 1083
|
|
|
QY 901 CAGCTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGTGAGAGCCCAAT 960
|
|
|
Db 1084 CAGCTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGTGAGAGCCCAAT 1143
|
|
|
QY 961 GAACTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAGAGAGAGCGCCCTCACTCCACCC 1020
|
|
|
Db 1144 GAACTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAGAGAGAGCGCCCTCACTCCACCC 1203
|
|
|
QY 1021 TGTGCGACTCTGTCTCAGAGGGAGAGCAAGAGAGCGCTGTGATCCCGCAGCGTG 1080
|
|
|
Db 1204 TGTGCGACTCTGTCTCAGAGGGAGAGCAAGAGAGCGCTGTGATCCCGCAGCGTG 1263
|
|
|
QY 1081 CCCAGCTGCGCCAGCGTGCAGCGCTGTGTGAGAGCAAGCGCTGTGTACAGGCGCTC 1140
|
|
|
Db 1264 CCCAGCTGCGCCAGCGTGCAGCGCTGTGTGAGAGCAAGCGCTGTGTACAGGCGCTC 1323
|
|
|
QY 1141 AGTGGCTGCACTGTGCGCAGACAGAGTGGAAAGAGAGATTAAGTCAAGGCTTCCTC 1200
|
|
|
Db 1324 AGTGGCTGCACTGTGCGCAGACAGAGTGGAAAGAGAGATTAAGTCAAGGCTTCCTC 1383
|
|
|
QY 1201 TCTCTGATATCAATCAAGTTTCAATTCAGCCAGCATGCGACATCTTACATGAGCTTC 1260
|
|
|
Db 1384 TCTCTGATATCAATCAAGTTTCAATTCAGCCAGCATGCGACATCTTACATGAGCTTC 1443
|
|
|
QY 1261 TCTCTGATCAAGATGCTTGTGAATCTTCAAACTTCCACTACCTGTGATGGAGCAAC 1320
|
|
|
Db 1444 TCTCTGATCAAGATGCTTGTGAATCTTCAAACTTCCACTACCTGTGATGGAGCAAC 1503
|
|
|
QY 1321 AAGCTATGCGAGTTTCCCTGTTCAGAACTATCGAGAGCAGATCCCGAAACAGAGTCTT 1380
|
|
|
Db 1504 AAGCTATGCGAGTTTCCCTGTTCAGAACTATCGAGAGCAGATCCCGAAACAGAGTCTT 1563
|
|
|
QY 1381 GATTAAGAGAGAGCAGCATCTCCCAAGAAAGCTGACAGCGCCAGCTTTCAGACAGCCAG 1440
|
|
|
Db 1564 GATTAAGAGAGAGCAGCATCTCCCAAGAAAGCTGACAGCGCCAGCTTTCAGACAGCCAG 1623
|
|
|

QY 1441 AGCAGCAGATTGATTCGGTTCAGAAACAGACAGAGTGGACCGCCAGAGTCCCTTTTA 1500
|
|
|
Db 1624 AGCAGCAGATTGATTCGGTTCAGAAACAGACAGAGTGGACCGCCAGAGTCCCTTTTA 1683
|
|
|
QY 1501 TCTCCACTGATTCGAAGTGGAGCGCTGGAGCAATTAACACACAGACTTCTTTTCGCG 1560
|
|
|
Db 1684 TCTCCACTGATTCGAAGTGGAGCGCTGGAGCAATTAACACACAGACTTCTTTTCGCG 1743
|
|
|
QY 1561 CTTTCACAGCAGCAGACAGACTCAAGAGTGTGCTGCGCTGGAGCTTAAGGGCTGGAC 1620
|
|
|
Db 1744 CTTTCACAGCAGCAGACAGACTCAAGAGTGTGCTGCGCTGGAGCTTAAGGGCTGGAC 1803
|
|
|
QY 1621 TGGATATTTGGCCCCCAGACTCTTACCCCTTCTGACAGCAGCTGTATTTTGC 1680
|
|
|
Db 1804 TGGATATTTGGCCCCCAGACTCTTACCCCTTCTGACAGCAGCTGTATTTTGC 1863
|
|
|
QY 1681 ACAGAGTCTCAGACTTCTTACCTGSCCAGCACTTACAGAGAGAGAGAGAGAGTCTT 1740
|
|
|
Db 1864 ACAGAGTCTCAGACTTCTTACCTGSCCAGCACTTACAGAGAGAGAGAGAGAGTCTT 1923
|
|
|
QY 1741 GCCTACAGCTGAGCAGCAGCTGCCACTTTCGAGAGCAAGCTTATCTGTGCGAGCGG 1800
|
|
|
Db 1924 GCCTACAGCTGAGCAGCAGCTGCCACTTTCGAGAGCAAGCTTATCTGTGCGAGCGG 1983
|
|
|
QY 1801 CAGAGCCAGTGCAGAGCTGACTCGCGCGAGAGCTGCGCATGAAGAGAGCCCTTTGAA 1860
|
|
|
Db 1984 CAGAGCCAGTGCAGAGCTGACTCGCGCGAGAGCTGCGCATGAAGAGAGCCCTTTGAA 2043
|
|
|
QY 1861 AAGCAGTTTAAACGAGAGCTGCAATGGAATTTTGAAGAGACATCATGTGAGAAAC 1920
|
|
|
Db 2044 AAGCAGTTTAAACGAGAGCTGCAATGGAATTTTGAAGAGACATCATGTGAGAAAC 2103
|
|
|
QY 1921 AGCTCACGGGAAAGCTGGGAAAGTGGGAGTCACTTACCTTTTGGGAGAGATGAA 1980
|
|
|
Db 2104 AGCTCACGGGAAAGCTGGGAAAGTGGGAGTCACTTACCTTTTGGGAGAGATGAA 2163
|
|
|
QY 1981 ATCATTTGAGTCTCC 1995
|
|
|
Db 2164 ATCATTTGAGTCTCC 2178
|
|
|

RESULT 13
AX180875 2732 bp DNA linear PAT 06-AUG-2001
LOCUS AX180875
DEFINITION Sequence 2 from Patent WO0146394.
ACCESSION AX180875
VERSION AX180875.1 GI:15132703
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Plozman, G.D., Martinez, R., Whyte, D., Manning, G., Sudarsanam, S.,
Hill, R.U., and Flanagan, P.
TITLE Mammalian protein phosphatases
JOURNAL Patent: WO 0146394-A 2 28-JUN-2001;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
source 1..2732
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 92.3%; Score 1842; DB 6; Length 2732;
Best local Similarity 99.8%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATTGGAATCTCAATTTGTTACTGAGAGTTGTGTGCTGTGAA 60
|
|
|
Db 538 ATGGCCCATGAGATGATTGGAATCTCAATTTGTTACTGAGAGTTGTGTGCTGTGAA 597
|
|
|

QY 61 AGTGAACGGAAAAAGTGTCTAATGATGAGCCGGCATTTGTGGAATACATATGATCC 120
| | | | |
Db 598 AGTGAACGGAAAAAGTGTCTAATGATGAGCCGGCATTTGTGGAATACATATGATCC 657
| | | | |
QY 121 CACATTTTGGAGCCATTAATATCACTGCTCCAGGTTATGAAGCGAAGTTGCAACAG 180
| | | | |
Db 658 CACATTTTGGAGCCATTAATATCACTGCTCCAGGTTATGAAGCGAAGTTGCAACAG 717
| | | | |
QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGATTCAGCCAAACATAAGGTTGATGAT 240
| | | | |
Db 718 GACAAAGTGTATTAACAGAGCTCATCCAGATTCAGCCAAACATAAGGTTGATGAT 777
| | | | |
QY 241 TGCAGTCAGAAAGGTGATTAAGATCAATCAAGCTCCAGAGATGTGCTCTCTCTCA 300
| | | | |
Db 778 TGCAGTCAGAAAGGTGATTAAGATCAATCAAGCTCCAGAGATGTGCTCTCTCTCA 837
| | | | |
QY 301 GACTGTTTCTCATCTGACTTCTGGGTAACTGAGAGAGAGCTTCACTGTTCACTG 360
| | | | |
Db 838 GACTGTTTCTCATCTGACTTCTGGGTAACTGAGAGAGAGCTTCACTGTTCACTG 897
| | | | |
QY 361 CTTCGAGGTGGGTTTGTGATGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
| | | | |
Db 898 CTTCGAGGTGGGTTTGTGATGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 957
| | | | |
QY 421 ACTCTAGTCCCTCACTGATTTCTCAGCCCTTGTGCTTACCTGCTTGAACATTTGGGCAACC 480
| | | | |
Db 958 ACTCTAGTCCCTCACTGATTTCTCAGCCCTTGTGCTTACCTGCTTGAACATTTGGGCAACC 1017
| | | | |
QY 481 CGAATTTTCCCAATCTTTATCTTGGCTGCGCAGAGATGCTCTCAACAGAGCTGATA 540
| | | | |
Db 1018 CGAATTTTCCCAATCTTTATCTTGGCTGCGCAGAGATGCTCTCAACAGAGCTGATA 1077
| | | | |
QY 541 CAGCAGATGGAGTTGGTTATGTGTTAATGCCAGTATACCTGTCCAAAGCCTGACTT 600
| | | | |
Db 1078 CAGCAGATGGAGTTGGTTATGTGTTAATGCCAGTATACCTGTCCAAAGCCTGACTT 1137
| | | | |
QY 601 ATCCCGAGTCTAATTTCCGCGGTGCGCTGTAATGAGAGCTTTGTGAAATTTTG 660
| | | | |
Db 1138 ATCCCGAGTCTAATTTCCGCGGTGCGCTGTAATGAGAGCTTTGTGAAATTTTG 1197
| | | | |
QY 661 CCGTGGTTGGCAAAATCAGTAGATTTCAATGAGAAAGCAAAAGCCTCAATGATGTGT 720
| | | | |
Db 1198 CCGTGGTTGGCAAAATCAGTAGATTTCAATGAGAAAGCAAAAGCCTCAATGATGTGT 1257
| | | | |
QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCGGCCAATGCTATGCTATCATGATG 780
| | | | |
Db 1258 CTAGTCACTGTTTAACTGGGATCTCCGCTCGGCCAATGCTATGCTATCATGATG 1317
| | | | |
QY 781 AAGAGATGAGCATGTCTTATGATGAACTTACAGATTTGTGAAAGAAAGAAAGCCTACT 840
| | | | |
Db 1318 AAGAGATGAGCATGTCTTATGATGAACTTACAGATTTGTGAAAGAAAGAAAGCCTACT 1377
| | | | |
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAAACTCTGGAATGAGAGAAAGTTAAGAAC 900
| | | | |
Db 1378 ATATCTCCAAACTTCAATTTTCTGGGCAAACTCTGGAATGAGAGAAAGTTAAGAAC 1437
| | | | |
QY 901 CAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGCACTTGAAGAGCCAAAT 960
| | | | |
Db 1438 CAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGCACTTGAAGAGCCAAAT 1497
| | | | |
QY 961 GAACTGTCTCTGCTCTCTCAGAGGTTGACAGAAAGAGAGAGCCCTCAATCCACCC 1020
| | | | |
Db 1498 GAACTGTCTCTGCTCTCTCAGAGGTTGACAGAAAGAGAGAGCCCTCAATCCACCC 1557
| | | | |
QY 1021 TGTGCGCACTCTGCTACTCAGAGGAGCAGAGCAAAAGCCCTGATCCCGCACGCTG 1080
| | | | |
Db 1558 TGTGCGCACTCTGCTACTCAGAGGAGCAGAGCAAAAGCCCTGATCCCGCACGCTG 1617
| | | | |
QY 1081 CCCAGCGTGGCCAGCGTGCAGCCGTGCTGTATGAGAGAGAGCCCGCTGTGATCAGCGGCTC 1140
| | | | |
Db 1618 CCCAGCGTGGCCAGCGTGCAGCCGTGCTGTATGAGAGAGAGCCCGCTGTGATCAGCGGCTC 1677
| | | | |
QY 1141 AGTGGGCTGCACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGGCTTCTTC 1200
| | | | |

Db 1678 AGTGGGCTGCACCTGTCCGACAGAGGCTGGAAGACAGCAATTAAGCTCAAGGTTCTTC 1737
| | | | |
QY 1201 TCTCTGATATCAATCACTTTATATTACGCCAGATGCGAGCATCTTACATGCTTC 1260
| | | | |
Db 1738 TCTCTGATATCAATCACTTTATATTACGCCAGATGCGAGCATCTTACATGCTTC 1797
| | | | |
QY 1261 TCCCTACAGAAAGTCTTTGGAATATCAAACTTCACTACTGTGATGGAGCAAC 1320
| | | | |
Db 1798 TCCCTACAGAAAGTCTTTGGAATATCAAACTTCACTACTGTGATGGAGCAAC 1857
| | | | |
QY 1321 AAGCTATGCAAGTCTCCCTGTTGAGAACTATCGAGACAGACTCCGAAACCACTCT 1380
| | | | |
Db 1858 AAGCTATGCAAGTCTCCCTGTTGAGAACTATCGAGACAGACTCCGAAACCACTCT 1917
| | | | |
QY 1381 GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGACAGCCGACCTTCAAGACCCAG 1440
| | | | |
Db 1918 GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGACAGCCGACCTTCAAGACCCAG 1977
| | | | |
QY 1441 AGCAAGCATTTGCTATTCGTTCAAGAACAGACAGAGTGGACCCGACAGTCCCTTTA 1500
| | | | |
Db 1978 AGCAAGCATTTGCTATTCGTTCAAGAACAGACAGAGTGGACCCGACAGTCCCTTTA 2037
| | | | |
QY 1501 TCTCCAGCTCATCGAAAGTGGAGAGCTGAGAGCAATTAACAGACAGCTTCTTTCGCG 1560
| | | | |
Db 2038 TCTCCAGCTCATCGAAAGTGGAGAGCTGAGAGCAATTAACAGACAGCTTCTTTCGCG 2097
| | | | |
QY 1561 CTTTCCACAGCAGCAGCAGCCTCAAGAGTGTGCTGAGCTGAGGCTTAAAGGCTGGCAC 1620
| | | | |
Db 2098 CTTTCCACAGCAGCAGCAGCCTCAAGAGTGTGCTGAGCTGAGGCTTAAAGGCTGGCAC 2157
| | | | |
QY 1621 TGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTACAGCAGCTGTATTTTGGC 1680
| | | | |
Db 2158 TGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTACAGCAGCTGTATTTTGGC 2217
| | | | |
QY 1681 ACAGAGTCTCACTTACTCTGCTCAGCAGCTTACAGAGAGAGAGAGAGCTTACTCT 1740
| | | | |
Db 2218 ACAGAGTCTCACTTACTCTGCTCAGCAGCTTACAGAGAGAGAGAGAGCTTACTCT 2277
| | | | |
QY 1741 GCTTACAGCTGAGCCAGCTGCCCCTTGGAGAGCAAGTATTTCTGTGCGCAGGCG 1800
| | | | |
Db 2278 GCTTACAGCTGAGCCAGCTGCCCCTTGGAGAGCAAGTATTTCTGTGCGCAGGCG 2337
| | | | |
QY 1801 CAGAGCCCAAGTGCAGAGCTGACTGCGCGGAGAGTGGCATGAGAGAGCCCTTTGAA 1860
| | | | |
Db 2338 CAGAGCCCAAGTGCAGAGCTGACTGCGCGGAGAGTGGCATGAGAGAGCCCTTTGAA 2397
| | | | |
QY 1861 AAGCAGTTTAAAGCAGAGAGCTGCCAAATGGAATTTGAGAGAGCATGTCAGAGAAC 1920
| | | | |
Db 2398 AAGCAGTTTAAAGCAGAGAGCTGCCAAATGGAATTTGAGAGAGCATGTCAGAGAAC 2457
| | | | |
QY 1921 AGTCAAGGGAAGAGCTGGGGAAGTGGGAGTCACTTCTTGGGCGAGCATGAA 1980
| | | | |
Db 2458 AGTCAAGGGAAGAGCTGGGGAAGTGGGAGTCACTTCTTGGGCGAGCATGAA 2517
| | | | |
QY 1981 ATCATTTAGGCTTCC 1995
| | | | |
Db 2518 ATCATTTAGGCTTCC 2532
| | | | |
RESULT 14
AX405700 3104 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 115 from Patent WO0222660.
DEFINITION AX405700
ACCESSION AX405700
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,

Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.
Novel nucleic acids and polypeptides
Patent: WO 0222660-A 115 21-MAR-2002;
HYSEQ, INC. (US)

FEATURES
source Location/Qualifiers

CDS
1..3104
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
159..2156
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD34807.1"
/db_xref="GI:2143840"
/translation="MAHEMCTOIYTERLVALLSESTREKULLIDSRPFVEYNTSHILE
AINTNSKIMKRRLOQDKVLTTELIOGSAKIKVIDSQKVVYDSQDVASLSDC
FLTYLGLKLEKSNVHLLAGFAEFSRCPEGLCEKSTLPTCIISQPCLPVANSIGPT
RILPNLYGQCDVANKELMOONGIYVLLAGLSRSATLAIAYIMKMDMLDEAYRPMK
ILPWLDKSDVPIERAKASNGCVLVHCLAGLSRSATLAIAYIMKMDMLDEAYRPMK
KRPTISPFNFIAGLDLYEKKIKNOTGASGPKSLKILHEKNEPVPALBSRGQSE
TLPSPCADSATSAGORPVHPASVPSVSVPSLDESPYQALSLADRLDE
SNKLRPSLDIKSVYSASMAASHFSSSEDALEYKPSITLIDGNKLCOPSPVOE
LSBOTPETSPEKPEASTPKLQIARPSDSQKRLHSVTSSTGTAQRLSLSPHRSQS
VEDNYHTSPLEGSTSQHLTKSAGLKGMSHDIILAPQTSLSLTSWTFATESHP
YASAIYGGASVAYSCSQLPTCGDQVYVRROKTSRQADSRGRSHEWSEPEKQFK
RRSCMEFHSIMSENRREBLKVGSGSFSGSMELIEVS"

ORIGIN

Query Match 92.3%; Score 1842; DB 6; Length 3104;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCCCATGATGATGATGAACTCAATGTTACTGAGAGGTGGTGGCTGCTGGGA 60
Db 159 ATGCCCATGATGATGATGAACTCAATGTTACTGAGAGGTGGTGGCTGCTGGGA 218
QY 61 AGTGAACGAAAAAGTGTCTAATGATAGCCGGCATTTGTGAAATACATATAC 120
Db 219 AGTGAACGAAAAAGTGTCTAATGATAGCCGGCATTTGTGAAATACATATAC 278
QY 121 CACATTTGGAAGCATTAATATCACTGCTCAAGTTATGAAGGAAAGTTGCAACG 180
Db 279 CACATTTGGAAGCATTAATATCACTGCTCAAGTTATGAAGGAAAGTTGCAACG 338
QY 181 GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATGAT 240
Db 339 GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATGAT 398
QY 241 TGCAGTCAGAAAGTGTGATTCAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
Db 399 TGCAGTCAGAAAGTGTGATTCAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 458
QY 301 GACTGTTTTCTCACTGTAATCTTGGGTAACCTGSAAGAGCTCAACTCTGTTCACTG 360
Db 459 GACTGTTTTCTCACTGTAATCTTGGGTAACCTGSAAGAGCTCAACTCTGTTCACTG 518
QY 361 CTTCGAGGTGGGTTTCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
Db 519 CTTCGAGGTGGGTTTCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 578
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGCTTACCTGTTGCCAATTTGGGCCAAC 480
Db 579 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGCTTACCTGTTGCCAATTTGGGCCAAC 638
QY 481 CGAATTTCTCCCAATCTTATCTTGGCTGCCAGCGAGATGCTCTCAAGAGAGTGTATA 540
Db 639 CGAATTTCTCCCAATCTTATCTTGGCTGCCAGCGAGATGCTCTCAAGAGAGTGTATA 698
QY 541 CAGCGAATGGGATTTGTTATGTTAAATGCGACCTATACCTGTCAAAAGCTGTACTT 600
Db 699 CAGCGAATGGGATTTGTTATGTTAAATGCGACCTATACCTGTCAAAAGCTGTACTT 758

QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGTAATGACAGCTTTGTGAGAAATTTTG 660
Db 759 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGTAATGACAGCTTTGTGAGAAATTTTG 818
QY 661 CCGTGTGGAAGAAACACTAATTTCAATTTGAGAAAGCAAAAGCTCCATGATGTGTT 720
Db 819 CCGTGTGGAAGAAACACTAATTTCAATTTGAGAAAGCAAAAGCTCCATGATGTGTT 878
QY 721 CTAGTGCACTGTTTATGCTGGATCTCCGCTCCGCAACATGCTATGCTCATCATG 780
Db 879 CTAGTGCACTGTTTATGCTGGATCTCCGCTCCGCAACATGCTATGCTCATCATG 938
QY 781 AAGAGATGGAATGCTCTTTAATGATGATTAAGATTGTAAGAAAGAAAGACTTACT 840
Db 939 AAGAGATGGAATGCTCTTTAATGATGATTAAGATTGTAAGAAAGAAAGACTTACT 998
QY 841 ATATCTCCAACTTCAATTTTGTGGGCCAATCTCGGCACTATGGAAGAAATTAAGAC 900
Db 999 ATATCTCCAACTTCAATTTTGTGGGCCAATCTCGGCACTATGGAAGAAATTAAGAC 1058
QY 901 CAGCTGAGCATCAGGCGCAAGCAAGCAAGCTCAAGCTGTGCACTTGAGAAAGCAAT 960
Db 1059 CAGCTGAGCATCAGGCGCAAGCAAGCAAGCTCAAGCTGTGCACTTGAGAAAGCAAT 1118
QY 961 GAACCTGTCCCTGCTGTCTCAGAGGTTGACAGAAAGCGAGAGCCCTCATGTCACC 1020
Db 1119 GAACCTGTCCCTGCTGTCTCAGAGGTTGACAGAAAGCGAGAGCCCTCATGTCACC 1178
QY 1021 TGTGCGCATCTGCTACTCAGAGGCAAGCAAGCAAGGCGCGCATCCGCGCAAGGTTG 1080
Db 1179 TGTGCGCATCTGCTACTCAGAGGCAAGCAAGGCGCGCATCCGCGCAAGGTTG 1238
QY 1081 CCCAGCGTCCAGAGGTCAGCCCTCGCTTTAAGAGACAGCCGCTGTATCAGCGCTC 1140
Db 1239 CCCAGCGTCCAGAGGTCAGCCCTCGCTTTAAGAGACAGCCGCTGTATCAGCGCTC 1298
QY 1141 AGTGGGCTGCACTGTCCGCGAGACAGGCTGGAACAGCAATAGCTCAAGCTTCTTC 1200
Db 1299 AGTGGGCTGCACTGTCCGCGAGACAGGCTGGAACAGCAATAGCTCAAGCTTCTTC 1358
QY 1201 TCTCTGATATCAATCAATGATTTCAATTTCAAGCCAGATGAGCATCTTACATGCTTC 1260
Db 1359 TCTCTGATATCAATCAATGATTTCAATTTCAAGCCAGATGAGCATCTTACATGCTTC 1418
QY 1261 TCTCTATCAAGAAATGCTTTGAAATTAATCAAACTTCACTGATGATGAGCAAC 1320
Db 1419 TCTCTATCAAGAAATGCTTTGAAATTAATCAAACTTCACTGATGATGAGCAAC 1478
QY 1321 AAGCTATGCGAGTTCTCCCTGTTCAAGAACTATCGGAGAGACTCCGAAACAGATCCT 1380
Db 1479 AAGCTATGCGAGTTCTCCCTGTTCAAGAACTATCGGAGAGACTCCGAAACAGATCCT 1538
QY 1381 GATTAAGAGAAAGCAGCATCCCAAGAAAGCTGAGACGCGCAAGCTTCAAGACACAG 1440
Db 1539 GATTAAGAGAAAGCAGCATCCCAAGAAAGCTGAGACGCGCAAGCTTCAAGACACAG 1598
QY 1441 AGCAAGGATTTGATTCGCTGCAAGCAAGAGAGTGCACCGCCCAAGAGTCCCTTTA 1500
Db 1599 AGCAAGGATTTGATTCGCTGCAAGCAAGAGAGTGCACCGCCCAAGAGTCCCTTTA 1658
QY 1501 TCTCACTGATGAGAGTGGAGAGCTGAGAGCAATTAACAACCAAGCTTCTCTTTGGG 1560
Db 1659 TCTCACTGATGAGAGTGGAGAGCTGAGAGCAATTAACAACCAAGCTTCTCTTTGGG 1718
QY 1561 CTTTCAACAGCGACAGCACTTCAAGAACTGCTGAGCTGAGCTTAAAGGCTGGCAC 1620
Db 1719 CTTTCAACAGCGACAGCACTTCAAGAACTGCTGAGCTGAGCTTAAAGGCTGGCAC 1778
QY 1621 TCGGATATCTGGGCCCCCAAGACTTACCTTCCCTGACAGAGCTGTATTTTGGC 1680
Db 1779 TCGGATATCTGGGCCCCCAAGACTTCTTACCTTCCCTGACAGAGCTGTATTTTGGC 1838
QY 1681 ACAGAGTCTCAACTTCTACTGCTCAGCACTTACGAGAGAGTGCAGTTACTCT 1740

Db 1839 ACGAGATCCCTCACTTCTACTGCTCAGCATCTAGGAGGAGCTGCACTTCTCT 1898
Qy 1741 GCTTACAGCTGCAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGGCAGCG 1800
Db 1899 GCTTACAGCTGCAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGGCAGCG 1958
Qy 1801 CAGAACCCAAAGTACAGAGCTGCTGCGCGGAGCTGGCATGAAGAGCCCTTTGA 1860
Db 1959 CAGAACCCAAAGTACAGAGCTGCTGCGCGGAGCTGGCATGAAGAGCCCTTTGA 2018
Qy 1861 AAGCAGTTTAAAGCAGAAAGCTCCAAATGGAATTTGAGAGAGCATCATGTGAGAGAA 1920
Db 2019 AAGCAGTTTAAAGCAGAAAGCTCCAAATGGAATTTGAGAGAGCATCATGTGAGAGAA 2078
Qy 1921 AGGTACCGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTCTTTGGGAGCATGAA 1980
Db 2079 AGGTACCGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTCTTTGGGAGCATGAA 2138
Qy 1981 ATCATTGAGCTTCC 1995
Db 2139 ATCATTGAGCTTCC 2153

RESULT 15
AX713989
LOCUS AX713989 2102 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 673 from Patent EP1293569.
ACCESSION AX713989
VERSION AX713989.1 GI:29888917
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Makamatsu,A., Sato,H., Ishii,S., Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R., Tamehika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahara,K. and Masuno,Y.
Full-length cDNAs
Patent: EP 1293569-A 673 19-MAR-2003;
Helix Research Institute (JP) ; Research Association for Biotechnology (JP)
Location/Qualifiers
source 1..2102
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 89.8%; Score 1791; DB 6; Length 2102;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

TITLE
JOURNAL
FEATURES
source 1..2102
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Qy 1 ATGGCCCATGAGATGATGGAATCAATTTGTAAGAGTGGTGGTCTGTCTGGAA 60
Db 56 ATGGCCCATGAGATGATGGAATCAATTTGTAAGAGTGGTGGTCTGTCTGGAA 115
Qy 61 AGTGAACGGAAAGTGTCTAATTGATAGCGCGCATTTGGAATCAATATACATCC 120
Db 116 AGTGAACGGAAAGTGTCTAATTGATAGCGCGCATTTGGAATCAATATACATCC 175
Qy 121 CACATTTTGGAGCCATTAATCACTGCTCCAGCTTATGAAGGAGTGGCAACAG 180
Db 176 CACATTTTGGAGCCATTAATCACTGCTCCAGCTTATGAAGGAGTGGCAACAG 235
Qy 181 GACAAAGTGTAAATTAAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATTGAT 240
Db 236 GACAAAGTGTAAATTAAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATTGAT 295
Qy 241 TGCAGTCAGAAAGTGTGATTTAGATCAAAAGCTCCCAAGATGTGCTCTCTCTCA 300

Db 296 TGCAGTCAGAAAGTGTGATTTAGATCAAAAGCTCCCAAGATGTGCTCTCTCTCA 355
Qy 301 GACTGTTTCTACATGTAATTTCTGGGTAACTGGAGAAAGAGTTCAACTCTTACCTG 360
Db 356 GACTGTTTCTACATGTAATTTCTGGGTAACTGGAGAAAGAGTTCAACTCTTACCTG 415
Qy 361 CTTCAGAGTGGTGTGATGTTCTCTCGTTTTCCTGGGCTCTGTGAAGAAATTC 420
Db 416 CTTCAGAGTGGTGTGATGTTCTCTCGTTTTCCTGGGCTCTGTGAAGAAATTC 475
Qy 421 ACTTATGTCCTACCTGATTTCTAGACCTTGTCTTACTGTTGCCAAATTTGGCCAA 480
Db 476 ACTTATGTCCTACCTGATTTCTAGACCTTGTCTTACTGTTGCCAAATTTGGCCAA 535
Qy 481 CGAATTTCTCCAAATCTTTATCTTGGCTGCCAGAGATGTCTCTCAAGAGAGCTGA 540
Db 536 CTAAATCTTCCAAATCTTTATCTTGGCTGCCAGAGAGTCTCTCAAGAGAGCTGA 595
Qy 541 CAGCAAAATGGGATTTGTTATGTTAAATGACAGTATACCTGTCCAAAGCTGACTT 600
Db 596 CAGCAAAATGGGATTTGTTATGTTAAATGACAGTATACCTGTCCAAAGCTGACTT 655
Qy 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTG 660
Db 656 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTG 715
Qy 661 CCGTGGTTGCAAAATCAGTATGTTTCAATTGAGAAAGCAAAAGCTTCCATGATGTT 720
Db 716 CCGTGGTTGCAAAATCAGTATGTTTCAATTGAGAAAGCAAAAGCTTCCATGATGTT 775
Qy 721 CTATGTCATGTTTATGCTGGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 780
Db 776 CTATGTCATGTTTATGCTGGGATCTCCGCTCCGACCATGCTTATGCTTACATCATG 835
Qy 781 AAGAGATGAGATGCTTTAGATGAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
Db 836 AAGAGATGAGATGCTTTAGATGAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 895
Qy 841 ATATCTCCAAATCTTCAATTTTCTGGGCCAACTCTGTGATGAGAAAGATTAAGAAC 900
Db 896 ATATCTCCAAATCTTCAATTTTCTGGGCCAACTCTGTGATGAGAAAGATTAAGAAC 955
Qy 901 CAGACTGAGCATCAGGCGCAAGCAAAAGCAAACTCAAGCTGCTGACCTGAGAAAGCA 960
Db 956 CAGACTGAGCATCAGGCGCAAGCAAAAGCAAACTCAAGCTGCTGACCTGAGAAAGCA 1015
Qy 961 GAACTGTCTCTGTCTCTCAGAGGTGAGCAGAAAGCGAGAGCGCCCTCACTGATCC 1020
Db 1016 GAACTGTCTCTGTCTCTCAGAGGTGAGCAGAAAGCGAGAGCGCCCTCACTGATCC 1075
Qy 1021 TGTGCCGACTGTGCTTACTCAGAGGAGCAGAGCAAAAGCGGTGATCCCGCAGCGT 1080
Db 1076 TGTGCCGACTGTGCTTACTCAGAGGAGCAGAGCAAAAGCGGTGATCCCGCAGCGT 1135
Qy 1081 CCCAGGTGCCAGAGGTGAGCGCTGCTGTGAGAGAGAGCGCGGTGTAACAGGCGTC 1140
Db 1136 CCCAGGTGCCAGAGGTGAGCGCTGCTGTGAGAGAGAGCGCGGTGTAACAGGCGTC 1195
Qy 1141 AGTGGCTCACCCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGGCTTCT 1200
Db 1196 AGTGGCTCACCCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGGCTTCT 1255
Qy 1201 TCTCTGATATCAATCAATCAATTTATTTAGCCAGAGATGAGCATCTTACATGCTTC 1260
Db 1256 TCTCTGATATCAATCAATCAATTTATTTAGCCAGAGATGAGCATCTTACATGCTTC 1315
Qy 1261 TCTCTATGAGAGATGCTTTGGAATTAATCAAACTTCCATCACTGTGATGGAGCAAC 1320
Db 1316 TCTCTATGAGAGATGCTTTGGAATTAATCAAACTTCCATCACTGTGATGGAGCAAC 1375
Qy 1321 AAGCTATGCAATTTCTCCCTGTTCAGAAATCTATGAGAGCAGATCCGAAACCACTCT 1380
Db 1376 AAGCTATGCAATTTCTCCCTGTTCAGAAATCTATGAGAGCAGATCCGAAACCACTCT 1435

```
QY 1381 GATTAAGGAGGAGCCAGCATCCCAAGAGCTGCAGACGGCCCTTCAGACAGCCAG 1440
    |||||
Db 1436 GATTAAGGAGGAGCCAGCATCCCAAGAGAGCTGCAGACTGCCAGCCCTTCAGACAGCCAG 1495
    |||||
QY 1441 AGCAAGGATTTGATTTGGATCAGAACCCAGCAGCAGTGGCACCGCCCAAGAGTCCCTTTTA 1500
    |||||
Db 1496 AGCAAGGATTTGATTTGGATCAGAACCCAGCAGCAGTGGCACCGCCCAAGAGTCCCTTTTA 1555
    |||||
QY 1501 TCTCCACTGCATGGAATGGAGAGCTGGAGAGCAATTACCAACAGCTTCTTTTGGGC 1560
    |||||
Db 1556 TCTCCACTGCATGGAATGGAGAGCTGGAGAGCAATTACCAACAGCTTCTTTTGGGC 1615
    |||||
QY 1561 CTTTCCACGAGCCAGCAGCAGCTCAAGAGTGTGCTGGCCCTGGCCCTTAAGGGCTGGCAC 1620
    |||||
Db 1616 CTTTCCACGAGCCAGCAGCAGCTCAAGAGTGTGCTGGCCCTTAAGGGCTGGCAC 1675
    |||||
QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGCAGTGTATTTTGGC 1680
    |||||
Db 1676 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGCAGTGTATTTTGGC 1735
    |||||
QY 1681 ACAGAGTCTTCACACTTCTACTCTGCTCAGCCATCTACGGAGGCACTGCCAGTTACTCT 1740
    |||||
Db 1736 ACAGAGTCTTCACACTTCTACTCTGCTCAGCCATCTACGGAGGCACTGCCAGTTACTCT 1795
    |||||
QY 1741 GCCTACAGCTGCAGCCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGGCAGGCGG 1800
    |||||
Db 1796 GCCTACAGCTGCAGCCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGGCAGGCGG 1855
    |||||
QY 1801 CAGAGGCCAAGTACAGAGCTGACTGCGGCGGAGCTGGCATGAGAGAGCCCTTTGAA 1860
    |||||
Db 1856 CAGAGGCCAAGTACAGAGCTGACTGCGGCGGAGCTGGCATGAGAGAGCCCTTTGAA 1915
    |||||
QY 1861 AAGCAGTTTAAAGCAGAGCTGCCAATGGAATTGGAGAGGCAATGTCAGAGAAC 1920
    |||||
Db 1916 AAGCAGTTTAAAGCAGAGCTGCCAATGGAATTGGAGAGGCAATGTCAGAGAAC 1975
    |||||
QY 1921 AGGTCAAGGGAAGAGCTGGGGAAGTGGGAGTCACTTAGCTTTTCGGGAGCATGAA 1980
    |||||
Db 1976 AGGTCAAGGGAAGAGCTGGGGAAGTGGGAGTCACTTAGCTTTTCGGGAGCATGAA 2035
    |||||
QY 1981 ATCATTGAGGTCTCC 1995
    |||||
Db 2036 ATCATTGAGGTCTCC 2050
    |||||
```

Search completed: June 22, 2004, 07:13:19
Job time : 12419 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2004, 02:52:01 ; Search time 7674 Seconds
(without alignments)
7763.233 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccatgagatgatg9.....tggaaatcatgaggtctcc 1995

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 200554

Minimum DB seq length: 1419

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_pig:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	2.7	3325	11 AK035652	AK035652 Mus muscu
2	39	2.0	2573	11 AK014441	AK014441 Mus muscu
3	19	1.0	1428	29 AY412468	AY412468 Pan trogl
4	19	1.0	1449	29 AY412467	AY412467 Homo sapi

C	5	19	1.0	1453	11 AK036059	AK036059 Mus muscu
C	6	19	1.0	2013	29 AY418586	AY418586 Mus muscu
C	7	19	1.0	2207	11 BC038231	BC038231 Homo sapi
C	8	19	1.0	2289	29 AY410969	AY410969 Mus muscu
C	9	19	1.0	2391	11 AK014482	AK014482 Mus muscu
C	10	19	1.0	2985	11 AK040537	AK040537 Mus muscu
C	11	18	0.9	1447	12 BM926476	BM926476 AGENCOURT
C	12	18	0.9	1506	12 BG118105	BG118105 602349869
C	13	18	0.9	1545	28 CC206026	CC206026 CH261-191
C	14	18	0.9	1564	12 BG757312	BG757312 602715303
C	15	18	0.9	2647	11 AK086250	AK086250 Mus muscu
C	16	18	0.9	2742	11 AK047025	AK047025 Mus muscu
C	17	18	0.9	3109	11 AK037848	AK037848 Mus muscu
C	18	18	0.9	3185	11 AK049870	AK049870 Mus muscu
C	19	18	0.9	3235	11 AK049360	AK049360 Mus muscu
C	20	18	0.9	3386	11 AK035947	AK035947 Mus muscu
C	21	18	0.9	3813	29 AY412149	AY412149 Homo sapi
C	22	18	0.9	3813	29 AY412150	AY412150 Pan trogl
C	23	18	0.9	4286	11 AK045549	AK045549 Mus muscu
C	24	18	0.9	5769	29 AY412182	AY412182 Homo sapi
C	25	18	0.9	5769	29 AY412183	AY412183 Pan trogl
C	26	17	0.9	6013	11 BC046934	BC046934 Homo sapi
C	27	17	0.9	1445	10 BE422076	BE422076 HMM019CF
C	28	17	0.9	1450	28 CC236993	CC236993 CH261-191
C	29	17	0.9	1454	11 BC000940	BC000940 Homo sapi
C	30	17	0.9	1459	29 AY401655	AY401655 Homo sapi
C	31	17	0.9	1471	12 B0054048	B0054048 AGENCOURT
C	32	17	0.9	1496	28 CC237048	CC237048 CH261-191
C	33	17	0.9	1502	8 BC012770	BC012770 Homo sapi
C	34	17	0.9	1504	29 AY411641	AY411641 Mus muscu
C	35	17	0.9	1528	11 AK017629	AK017629 Mus muscu
C	36	17	0.9	1580	11 AK033751	AK033751 Mus muscu
C	37	17	0.9	1587	28 CC189276	CC189276 CH261-17J
C	38	17	0.9	1589	29 AY403911	AY403911 Homo sapi
C	39	17	0.9	1593	14 CF110912	CF110912 Shultzom
C	40	17	0.9	1607	12 BG566227	BG566227 602585083
C	41	17	0.9	1614	29 AG158840	AG158840 Pan trogl
C	42	17	0.9	1618	11 AK075981	AK075981 Mus muscu
C	43	17	0.9	1635	29 AY402679	AY402679 Pan trogl
C	44	17	0.9	1640	12 B1552587	B1552587 603197914
C	45	17	0.9	1683	11 AK078482	AK078482 Mus muscu

ALIGNMENTS

RESULT 1	AK035652	3325 bp	mRNA	linear	HTC 19-SEP-2003
LOCUS	AK035652				
DEFINITION	Mus musculus adult male urinary bladder cDNA, RIKEN full-length enriched library, clone:9530081P05 product:TRUNCATED MRPK PHOSPHATASE 7 homolog [Homo sapiens], full insert sequence.				
ACCESSION	AK035652	GI:26330815			
VERSION	AK035652.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	Carinci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meeth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
AUTHORS	Carinci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				

REFERENCE
AUTHORS
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Komoto,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
PUBMED
20530913
11076861

REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL
AUTHORS
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
REFERENCE
AUTHORS
6
(bases 1 to 3325)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoch,H., Kawai,J., Kojima,Y., Kondo,S., Komoto,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Takeda,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toyota,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

JOURNAL
COMMENT
CNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 3325
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:9530081F05"
/db_xref="MGI:2359973"
/db_xref="taxon:10090"
/clone="9530081F05"
/sex="male"
/tissue_type="urinary bladder"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
460..1080
/note="unnamed protein product; TRUNCATED MAPK PHOSPHATASE
7 homolog [Homo sapiens] (SPTK1) [Q960S2, evidence:FASTV,
96.7%id, 86%length, match=369]
putative"
/codon_start=1
/protein_id="BAC29138.1"

FEATURES
source

ORIGIN
Query Match 2.7%; Score 53; DB 11; Length 3325;
Best Local Similarity 100.0%; Pred. No. 4,2e-15;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
AK014441
LOCUS
DEFINITION
Mus musculus 18 days pregnant adult female placenta and extra
embryonic tissue cDNA, RIKEN full-length enriched library,
clone:3830417M17 product:TRUNCATED MAPK PHOSPHATASE 7 homolog [Homo
sapiens], full insert sequence.
ACCESSION
AK014441 GI:12852294
VERSION
AK014441.1 GI:12852294
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL
MEDLINE
PUBMED
99279253
10349636

REFERENCE
AUTHORS
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Komoto,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL
MEDLINE
PUBMED
20499374
11042159

REFERENCE
AUTHORS
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Komoto,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
PUBMED
20530913
11076861

REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL
AUTHORS
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
REFERENCE
AUTHORS
6
(bases 1 to 2573)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M.,
Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
Hirooka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,

Kasuhara, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koyama, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Onno, M., Okazaki, Y., Okita, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Togabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-1-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome-gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

For further details, our web site (<http://genome-gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGGAGGAGGAGATCCAAAGGACCTCTTTTCTTTTCTTNN 3']. cDNA was prepared by using trisphosphate thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction for Rot = 50.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGGAGGAGATCTCGAGTTAATTAATATGCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified plasmidscript KS(+) after bulk excision from lambda damB Flc I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B.

```

REFERENCE      1 (bases 1 to 1428)
AUTHORS        Clark,A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
               Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
               Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sznitsky, J.J.,
               Adams, M.D. and Cargill, M.
TITLE          Inferring nonneutral evolution from human-chimp-mouse orthologous
               gene trios
JOURNAL        Science 302 (5652), 1960-1963 (2003)
PUBMED         14671302
REFERENCE      2 (bases 1 to 1428)
AUTHORS        Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
               Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
               Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sznitsky, J.J.,
               Adams, M.D. and Cargill, M.
TITLE          Direct Submision
COMMENT        Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
               Rockville, MD 20850, USA
FEATURES       This sequence was made by sequencing genomic exons and ordering
               them based on alignment.
SOURCE         Location/Qualifiers
               1..1428
               /organism="Pan troglodytes"
               /mol_type="genomic DNA"
               /db_xref="taxon:9598"
               <1..>1428
               /gene="DUSP10"
               /locus_tag="HGM4567"
ORIGIN

```

ORIGIN

Query Match	1.0%;	Score 19;	DB 29;	length 1428;
Best Local Similarity	100.0%;	Pred. No. 1.1e+03;		
Matches	19;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

QY	745	TCCCGCTCCGCCACCATCG	763
Db	1237	TCCCGCTCCGCCACCATCG	1255

```

/db_xref="PANTOM DB:3830417M17"
/db_xref="MGI:1903116"
/db_xref="Taxon:10090"
/clone="3830417M17"
/sex="female"
/tissue_type="Placenta and extra embryonic tissue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18 days pregnant adult"
1. 2573
/ncbi="TRUNCATED MAPK PHOSPHATASE 7 homolog (Homo sapiens)
(SPR|096082, evidence: FASRY, 96.7%ID, 86%length,
match=369)"

misc_feature

ORIGIN

Query Match
Beat Local Similarity 100.0%; Score 39; DB 11; Length 2573;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1869 TAAACGCAAGAGCTCCCAATGGAATTTGGAGAGGCAAT 1907
|||||
40 TAAACGCAAGAGCTCCCAATGGAATTTGGAGAGGCAAT 78

RESULT 3
AY412468 1428 bp DNA linear GSS 16-DEC-2003
LOCUS AY412468
DEFINITION Pan troglodytes DUSP10 gene, VIRtual TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY412468
VERSION AY412468.1 GI:39768433
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

```

LOCUS	AV412467	1449 bp	DNA	linear	GSS 16-DEC-2003
DEFINITION	Homo sapiens DUSP10 gene, VIRUTAL TRANSCRIPT, partial sequence.				
ACCESSION	AV412467				
VERSION	AV412467.1	GI:39768432			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1449)				
AUTHORS	Clark,A.G., Glatowski,S., Nielson,R., Thomas,P., Kejalrajal,A., Todd,M.A., Tanendbaum,D.M., Clivello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 1449)				
AUTHORS	Clark,A.G., Glatowski,S., Nielson,R., Thomas,P., Kejalrajal,A., Todd,M.A., Tanendbaum,D.M., Clivello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	Location/Qualifiers				
SOURCE	1..1449				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				

```

/db_xref="taxon:9606"
<1..>1449
gene

```

OPTIMUM

```
/gene="DUSP10"  
/locus_tag="HCM4567"
```

Query Match	1.0%;	Score 19;	DB 29;	Length 1449;
Best Local Similarity	100.0%;	Pred. No. 1.1e+03;		
Matches	19;	Conservative	0;	Mismatches 0;
				Indels

DY	745 TCCGCTCGGCACCATCG	763
Db	1237 TCCGCTCGGCACCATCG	1255

RESULT 5	AK036059/c	LOCUS	AK036059	1453 bp	mRNA	linear	HTC 19-SEP-2003
DEFINITION	Mus musculus 16 days neonate cerebellum cDNA						

ACCESSION	AK036059
VERSION	AK036059.1
KEYWORDS	HTC, CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

REFERENCE
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
2

TITLE	JOURNAL	PUBLISHED	REFERENCES
Normalization and subtraction of cap-trapper-selected CNAs to prepare full-length cDNA libraries for rapid discovery of new genes	Genome Res.	10 (10), 1617-1630 (2000)	20499374
	PUBLISHED	11042155	
		3	
			Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, Y., Kawamura, T.

Matsumoto, R., Matsumoto, R., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, K., Izawa, M., Ohara, E., Wataniki, M., Onoda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Kameda, Y., Muramatsu, M., Inoue, Y., Kiz, A., and Hayashizaki, Y. TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913

AUTHORS	TITLE	JOURNAL	EXPERIENCE	AUTHORS
4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	Functional annotation of a full-length mouse cDNA collection	Nature 409, 685-690 (2001)	5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase II Team				

JOURNAL
OF 60, 770 Full-length cDNAs
REFERENCE
Nature 420, 563-573 (2002)
AUTHORS
6 (bases 1 to 1453)
Aadachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagata, T., Hara, A., Hashizume, M.,
Hayashida, S., Hayatsu, N., Higashimoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imocani, K., Ishii, Y., Itochi, N., Kagawa, I., Kasukabe, T.,
Kato, H., Kawaji, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.

TITLE
JOURNAL

COMMENT

FEETIMES

CDS

ORIGIN

Query Match	1.0%;	Score 19;	DB 11;	Length 1453;
Best Local Similarity	100.0%;	Pred. No. 1.1e+03;		
Matches 19;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

```

      1 / 45 ACAGCTGCAGCCAGCTGCC 176
            |||||
Db       737 ACAGCTGCAGCCAGCTGCC 719

```

RESULT 6	AY418586	2013 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	AY418586				
DEFINITION	Mus musculus HCM6606 gene, VIRAL TRANSCRIPT, partial sequence, genomic structure, Genbank				

REFERENCE
AUTHORS
1 (bases 1 to 2013)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,

TITLE Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2013)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shtinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES Location/Qualifiers
source
1..2013
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2013
/locus_tag="HGM6606"
ORIGIN
gene
Query Match 1.0%; Score 19; DB 29; Length 2013;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1077 CGTGCCAGCGCTGCCGAGC 1095
|||||
Db 1132 CGTGCCAGCGCTGCCGAGC 1150
|||||
RESULT 7 BC038231 2207 bp mRNA linear HTC 01-OCT-2002
LOCUS BC038231 Homo. sapiens, Similar to dual specificity phosphatase 8, clone
DEFINITION IMAGE:5547764, mRNA.
ACCESSION BC038231.1 GI:23398534
VERSION BC038231.1
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2207)
AUTHORS Strausberg, R.
JOURNAL Direct Submission
PUBMED Submitted (30-SEP-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
TITLE NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapsb-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gatherburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
ORIGIN Contact: nisc.mgc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legadi, R.,
Maduro, O.L., Mastello, C., Maskeri, B., Maserian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stancipop, S., Thomas, P.J., Touchman, J.W.,
Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LIML at: http://image.liml.gov
Series: IRAK Plate: 79 Row: a Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4758211
This clone has the following problem: frame shifted.
FEATURES Location/Qualifiers
source
1..2207
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5547764"
/issue_type="Skin, melanotic melanoma."
/clone_1ib="NIH MGC_72"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
ORIGIN
Query Match 1.0%; Score 19; DB 11; Length 2207;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 766 ATCGCTACATCATGAGA 784
|||||
Db 903 ATCGCTACATCATGAGA 921
|||||
RESULT 8 AY410969 2289 bp DNA linear GSS 12-DEC-2003
LOCUS AY410969 Mus musculus HGM4067 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY410969
VERSION AY410969.1 GI:39766937
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2289)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shtinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2289)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shtinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES Location/Qualifiers
source
1..2289
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2289
/locus_tag="HGM4067"
ORIGIN
gene
Query Match 1.0%; Score 19; DB 29; Length 2289;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1792 CGCAGCGCGCAGAGCCAA 1810
|||||
Db 246 CGCAGCGCGCAGAGCCAA 228
|||||

LOCUS	AK014482	2391 bp	mRNA	linear	HTC 20-SEP-2003
DEFINITION	Mus musculus 14 days embryo liver cDNA, RIKEN full-length enriched library, clone:4432409M07 product:hypothetical dDENN domain/DENN (AEK-3) domain containing protein, full insert sequence.				
ACCESSION	AK014482				
VERSION	AK014482.1				
KEYWORDS	HTC, CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
PUBMED	99279253				
REFERENCE	10349636				
AUTHORS	2				
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
PUBMED	20499374				
REFERENCE	11042159				
AUTHORS	3				
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, Y., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubiki, M., Yoneda, Y., Iihikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kiru, A. and Hayashizaki, Y.				
JOURNAL	RIKEN integrated sequence analysis (RISA) system-384-format				
PUBMED	sequencing pipeline with 384 multicapillary sequencer				
REFERENCE	Genome Res. 10 (11), 1757-1771 (2000)				
AUTHORS	4				
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
PUBMED	11076861				
REFERENCE	20530913				
AUTHORS	5				
REFERENCE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.				
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs				
JOURNAL	Nature 420, 563-573 (2002)				
PUBMED	6 (bases 1 to 2391)				
REFERENCE	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Aikawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hironaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Shibata, K., Shibata, Y., Shingana, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tajima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.				
JOURNAL	Direct Submission				
PUBMED	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222).				
REFERENCE	TITLE				
AUTHORS	JOURNAL				

COMMENT
Fax:81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGGACAGAAGATCCCAAGAGCTCTTTTCTTTTTTNN 3', cDNA was prepared by using tris(holo) thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10⁶ and subtraction to Rot = 100⁶. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGGACAGATCTCCGATTAAATTAATATCCCCCCCCC 3'. cDNA was cleaved with BamHI and XhoI. Vector: a modified phuescript KS(+) after bulk excision from lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES
source
1..2391
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="PANTOM:DB:4432409M07"
/db_xref="MG1:1903235"
/db_xref="taxon:10090"
/clone="4432409M07"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="14 days embryo"
110..913
/note="unnamed protein product; hypothetical dDENN domain/DENN (AEX-3) domain containing protein [Pfam] PF03455, InterPro|IPR001194, evidence: InterPro) putative"
/codon_start=1
/protein_id="BAB29384.1"
/db_xref="GI:12852367"
/translation="MGSTETRHPPAMDFEAGCCNSLEEDPILROFPDPFOEOELQWMPVRCFPEIDIREPPSPAVQHFTALDVLGNRRGFCRLAGASCICLSHFWEVFYKLINNVGLLONOVABAEELLQOHPLGFRSGSEMDSSITVRECCGILLPALNCISCFVAADSLPSIPENRYLTLLVYAVDENVIGLPAALLAEKRVILTAKLSIVRDHRSRGDRITKPFCSVPIPSEOSEVAQNLFRTGT"

ORIGIN
Query Match 1.0%; Score 19; DB 11; Length 2391;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Gy 1792 CGCAGCGGGCGCAGAGCCAA 1810
|||||||
Db 372 CGCAGCGGGCGCAGAGCCAA 354
|||||||

RESULT 10 AK040537/c
LOCUS AK040537/c
DEFINITION Mus musculus o day neonate thymus cDNA, RIKEN full-length enriched library, clone:A430105K15 product:hypothetical dDENN domain/DENN (AEX-3) domain containing protein, full insert sequence.
ACCESSION AK040537
VERSION AK040537.1 GI:26087915
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci P. and Hayashizaki Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PubMed 10349636

REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. <i>Genome Res.</i> 10 (10), 1617-1630 (2000)
JOURNAL		
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, T., Nishie, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer <i>Genome Res.</i> 10 (11), 1757-1771 (2000)
JOURNAL		
REFERENCE	4	20530913 11076861
JOURNAL		
REFERENCE	5	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection <i>Nature</i> 409, 685-690 (2001)
JOURNAL		
REFERENCE	6	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs <i>Nature</i> 420, 563-573 (2002)
JOURNAL		
REFERENCE	7	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fuyushida, K., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, P., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komoto, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohashi, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanabe, Y., Tatem, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-research@riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
JOURNAL		
REFERENCE	8	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers 1. 2985
JOURNAL		
REFERENCE	9	1. 2985 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:A430105K15" /db_xref="MGI:2405221" /db_xref="taxon:10090" /clone="A430105K15" /issue_type="thymus" /clone_type="RIKEN full-length enriched mouse cDNA library" /dev_stage="0 day neonate"
JOURNAL		

```

misc_feature      1..2985
                  /note="hypothetical dDENN domain/DENN (AEX-3) domain
                  containing protein (Pfam|PF03455, InterPro|IPR001194,
                  evidence: InterPro)"
ORIGIN
Query Match      1.0%; Score 19; DB 11; Length 2985;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY               1792 CGCAGCGCGCAGAGCCAA 1810
                  |||||
Db               368 CGCAGCGCGCAGAGCCAA 350

RESULT 11
LOCUS            BM926476                1447 bp    mRNA    linear    EST 12-MAR-2002
DEFINITION      AGNCNCURT_6644904 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5766863
ACCESSION       BM926476
VERSION         BM926476.1 GI:19376855
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS         1 (bases 1 to 1447)
TITLE           NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL         National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT         Unpublished (1999)
                  Contact: Robert Strausberg, Ph.D.
                  Email: cgabs-remail.nih.gov
                  Tissue Procurement: Life Technologies, Inc.
                  cDNA Library Preparation: Life Technologies, Inc.
                  DNA Sequencing by: Agencourt Bioscience Corporation (LNL)
                  Clone distribution: MGC Clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LNL at:
                  http://image.llnl.gov
                  Plate: LLM12824 row: C column: 24
                  High quality sequence, stop: 557.
                  Location/Qualifiers
                    1..1447
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5766863"
                      /lab_host="DH10B"
                      /clone_1b="NIH MGC 122"
                      /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
                      site.1: NotI; Site 2: EcoRV (destroyed); RNA source
                      anonymous pool of 24 week female lung, 16 week female
                      spleen, and 20-22 week male spleens. Library is oligo-dT
                      primed and directionally cloned (EcoRV site is destroyed
                      upon cloning). Average insert size 1.4 kb, insert size
                      range 1-3 kb. Library is normalized and enriched for
                      full-length clones and was constructed by C. Gruber
                      (Invitrogen). Research Genetics tracking code 026. Note:
                      this is a NIH_MGC Library."
ORIGIN
Query Match      0.9%; Score 18; DB 12; Length 1447;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY               819 TGTGAAGAGAAAAAGACC 836
                  |||||
Db               1236 TGTGAAGAGAAAAAGACC 1253

RESULT 12
RG118105/c

```

LOCUS BG118105 1506 bp mRNA linear EST 30-JAN-2001
 DEFINITION 602349869F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:444488 5',
 mRNA sequence.
 ACCESSION BG118105
 VERSION BG118105.1 GI:12611611
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1506)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapds-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LHAM10221 row: i column: 17
 High quality sequence stop: 2.
 Location/Qualifiers
 1..1506
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:444488"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC_90"
 /note="Organ: liver; Vector: pCMV-SORT6; Site_1: NctI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC library."

ORIGIN

Query Match 0.9%; Score 18; DB 12; Length 1506;
 Best Local Similarity 100.0%; Pred. No. 3.5e+03;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1547 GCTTCCTTTCCGCGCTTT 1564
 |||||
 DB 1506 GCTTCCTTTCCGCGCTTT 1489

RESULT 13
 CC206026 1545 bp DNA linear GSS 09-MAY-2003
 LOCUS CH261-191118_Sp6.1 CH261 Gallus gallus genomic clone CH261-191118,
 DEFINITION genomic survey sequence.
 ACCESSION CC206026
 VERSION CC206026.1 GI:30489567
 KEYWORDS GSS.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 1545)
 AUTHORS Krenitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
 Warren, W., Graves, T., Wards, E. and Wilson, R.
 TITLE Gallus gallus BAC End Reads
 JOURNAL Unpublished (2003)
 COMMENT Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@wustl.edu
 Insert Length: 182000 Std Error: 0.00

Seq primer: Sp6 ATTTAGTGACACTATAG
 Class: BAC ends
 High quality sequence start: 250
 High quality sequence stop: 900.
 Location/Qualifiers
 1..1545
 /organism="Gallus gallus"
 /mol_type="genomic DNA"
 /strain="Red Jungle Fowl"
 /db_xref="taxon:9031"
 /clone="CH261-191118"
 /sex="female"
 /cell_line="UCD001, inbred 256"
 /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
 CH261 Female Chicken library - for library and clone
 ordering information: <http://www.chori.org/bacpac>"

ORIGIN

Query Match 0.9%; Score 18; DB 28; Length 1545;
 Best Local Similarity 100.0%; Pred. No. 3.5e+03;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 813 CAGATTGTGAAGAAA 830
 |||||
 DB 816 CAGATTGTGAAGAAA 799

RESULT 14
 BG757312 1564 bp mRNA linear EST 15-MAY-2001
 LOCUS 602715303F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4855296 5',
 DEFINITION mRNA sequence.
 ACCESSION BG757312
 VERSION BG757312.1 GI:14067965
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1564)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapds-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LNCM1704 row: f column: 01
 High quality sequence stop: 272.
 Location/Qualifiers
 1..1564
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4855296"
 /tissue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

ORIGIN

Query Match 0.9%; Score 18; DB 12; Length 1564;

Best Local Similarity 100.0%; Pred. No. 3.6e+03;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1690 TCACACTTCTACTCTGCC 1707
 |||
 1276 TCACACTTCTACTCTGCC 1293

RESULT 15

AK086250/c
 LOCUS 2647 bp mRNA linear HTC 20-SEP-2003
 DEFINITION Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched

library, clone:ID930016F01 product:unknown EST, full insert

sequence.

ACCESSION AK086250
 VERSION AK086250.1 GI:26103338

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Komno, H., Akiyama, J., Nishi, K., Kitsuishi, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system—384-format
 sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 PANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

MEDLINE 11076861

PUBMED 11076861

REFERENCE

AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 11076861

PUBMED 11076861

REFERENCE

AUTHORS Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission Yoshihide Hayashizaki, The Institute of
 Subtitled (16-Apr-2002) Laboratory for Genome
 Exploration and Chemical Research (RIKEN), Laboratory for Genome
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suenryo-cho, Teurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

FEATURES
 source
 1.2647
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM:DB:D930016F01"
 /db_xref="MGI:2423638"
 /db_xref="taxon:10090"
 /clone="D930016F01"
 /issue_type="head"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="15 days embryo"
 1.2647
 /note="unknown EST (GB|BB749563, evidence: BLASTN, 99%,
 match=375)"

misc_feature

Query Match 0.9%; Score 18; DB 11; Length 2647;
 Best Local Similarity 100.0%; Pred. No. 4.2e+03;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1465 ACCAGCAGCAGTGCGACC 1482
 |||
 Db 2281 ACCAGCAGCAGTGCGACC 2264

Search completed: June 22, 2004, 09:21:26

Job time : 7680 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 18:35:51 ; Search time 1140 Seconds
(without alignments)
7434.347 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995

Sequence: 1 atgagccatgagatgattg9.....tggaaatcattgaggtctcc 1995

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 772728

Minimum DB seq length: 1419

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: geneseq119808:.*
2: geneseq119908:.*
3: geneseq120008:.*
4: geneseq12001a:.*
5: geneseq12001b:.*
6: geneseq12002a:.*
7: geneseq12003a:.*
8: geneseq12003b:.*
9: geneseq12003c:.*
10: geneseq12004a:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1995	100.0	5450	6 ACC60559	ACC60559 Polynucle
2	1995	100.0	5450	6 ACC60572	ACC60572 Polynucle
3	1893	94.9	3059	6 AAS15768	AAS15768 CDNA sequ
4	1893	94.9	3496	6 ABK47596	ABK47596 CDNA enco
5	1893	94.9	3544	5 AAS14639	AAS14639 Human CDN
6	1893	94.9	3544	6 ABK49402	ABK49402 CDNA enco
7	1893	94.9	3766	6 ABK14474	ABK14474 Human pro
8	1893	94.9	4790	6 ABN83966	ABN83966 Human gen
9	1893	94.9	5145	5 ABV20833	ABV20833 Human pro
10	1893	94.9	5145	5 ABV21080	ABV21080 Human pro
11	1893	94.9	5145	5 ABV26680	ABV26680 Human pro
12	1893	94.9	5145	5 ABV20978	ABV20978 Human pro
13	1893	94.9	5145	5 ABV21092	ABV21092 Human pro
14	1893	94.9	5145	5 ABV21312	ABV21312 Human pro
15	1893	94.9	5145	5 ABV21316	ABV21316 Human pro
16	1893	94.9	5145	5 ABV26826	ABV26826 Human pro
17	1893	94.9	5145	5 ABV27131	ABV27131 Human pro
18	1893	94.9	5145	5 ABV26923	ABV26923 Human pro
19	1893	94.9	5145	5 ABV27135	ABV27135 Human pro
20	1893	94.9	5145	5 ABV28657	ABV28657 Human pro
21	1893	94.9	5145	5 ABV22827	ABV22827 Human pro
22	1893	94.9	5145	5 ABV26934	ABV26934 Human pro
23	1849	92.7	2966	4 AAH99685	AAH99685 Human pro

24	1842	92.3	2732	4 AAD09492	AA009492 Human SGP
25	1842	92.3	3104	6 ABNS59704	ABNS59704 Novel hum
26	1791	89.8	2102	7 ADA53105	ADA53105 Human cod
27	1791	89.8	5111	6 ACC60521	ACC60521 Polynucle
28	1593	79.8	2118	4 AAF30479	AA130479 Human pro
29	1418	71.1	3332	6 ABK48378	ABK48378 CDNA enco
30	79	4.0	2756	6 ACC60560	ACC60560 Polynucle
31	21	1.1	2651	4 AAH54080	AAH54080 S. epider
32	21	1.1	4051	4 AAH54931	AAH54931 S. epider
33	20	1.0	1461	5 AAS72277	AAS72277 DNA enco
34	19	1.0	1422	6 ABLS8978	AB158978 HPV16-L2
35	19	1.0	1449	8 ADA02686	ADA02686 Human DUS
36	19	1.0	1449	9 ADB72424	ADB72424 Human DUS
37	19	1.0	1830	3 AAC64262	AAC64262 Human dua
38	19	1.0	1909	6 ABT07757	ABT07757 Breast ca
39	19	1.0	1909	6 ADA02685	ADA02685 Human DUS
40	19	1.0	1909	9 ADB72423	ADB72423 Human DUS
41	19	1.0	1917	9 ADE07547	Ade07547 Novel cod
42	19	1.0	2050	4 AAF63577	AA63577 Human pho
43	19	1.0	2081	4 AA199312	AA199312 Human exc
44	19	1.0	2081	5 AA163662	AA163662 Human kid
45	19	1.0	2170	2 AAV52445	AAV52445 Streptoco

ALIGNMENTS

RESULT 1	
ID ACC60559	standard; cdna; 5450 BP.
XX	
XX ACC60559;	
XX	
19-JUN-2003	(first entry)
XX	
DE Polynucleotide relating to the invention SEQ ID NO: 108.	
XX	
KW Gene, ser, antiproliferative; hepatotropic; nephrotropic; antiarthritic;	
KW antiproliferative; cardiatic; cytostatic; gene therapy; liver disease;	
KW proliferative disorder; renal failure; cardiovascular disorder;	
KW immunological disorder; arthritis; psoriasis; congenital heart defect;	
KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.	
XX	
OS Homo sapiens.	
XX	
PN WO200257460-A2.	
XX	
PD 25-JUL-2002.	
XX	
PF 20-DEC-2001; 2001MO-US050459.	
XX	
PR 20-DEC-2000; 2000US-0256868P.	
XX	
PR 30-MAR-2001; 2001US-0280186P.	
XX	
PR 01-MAY-2001; 2001US-0287735P.	
XX	
PR 05-JUN-2001; 2001US-0295848P.	
XX	
PR 25-JUN-2001; 2001US-0300465P.	
XX	
PA (BRIM) BRISTOL-MYERS SQUIBB CO.	
XX	
PI Jackson DG, Feder J, Nelson T, Muntier G, Ramanathan C, Lee L;	
XX	
PI Stewens N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;	
XX	
PI Kysceck S, Mcatee P, Suchard S, Banas D;	
XX	
DR WPI; 2002-559721/64.	
XX	
DR P-PADB; ABR52381.	
XX	
PT Novel polynucleotides encoding human phosphatase polypeptides useful in	
XX	
PT the prevention or treatment of e.g. proliferative and cardiovascular	
XX	
PS disorders.	
XX	
PS Claim 1, Fig 13; 801pp; English.	
XX	
CC The invention relates to a novel isolated nucleic acid comprising a	

CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antirhectic, antiproliferative, cardiant, and cytosolic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX

Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;

Query Match 100.0%; Score 1995; DB 6; Length 5450;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCCATGAGATGATGGAAGCAATGTTACTGAGAGTGGTGGCTGCTGGA 60
DB 538 ATGCCCATGAGATGATGGAAGCAATGTTACTGAGAGTGGTGGCTGCTGGA 597
QY 61 AGTGAACGAAAAAGTCTCTAATTGATAGCCGCAATTTGTGAATACAATACATCC 120
DB 598 AGTGAACGAAAAAGTCTCTAATTGATAGCCGCAATTTGTGAATACAATACATCC 657
QY 121 CACATTTTGAAGCATTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACAG 180
DB 658 CACATTTTGAAGCATTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACAG 717
QY 181 GACAAAGTTAATTAACAAGCTCATCAGCATTCAGCAAACTAAGTTGACATTAAT 240
DB 718 GACAAAGTTAATTAACAAGCTCATCAGCATTCAGCAAACTAAGTTGACATTAAT 777
QY 241 TGCAGTGAAGAGTTGATTAAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 300
DB 778 TGCAGTGAAGAGTTGATTAAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 837
QY 301 GACTGTTTCTCACTGACTTCTGGGTAACCTGGAAGAGCTCAACTCTGTTTCCAG 360
DB 838 GACTGTTTCTCACTGACTTCTGGGTAACCTGGAAGAGCTCAACTCTGTTTCCAG 897
QY 361 CTTGCAAGTGGTGTGCTGAGTTCTCTGTTTCCCTGACCTCTGTAAGAGAAATCC 420
DB 898 CTTGCAAGTGGTGTGCTGAGTTCTCTGTTTCCCTGACCTCTGTAAGAGAAATCC 957
QY 421 ACTCTAGTCCCTAAGCTGATTTCTCAGCCCTGTTACCTGTTCCCAACTTTGGCCAC 480
DB 958 ACTCTAGTCCCTAAGCTGATTTCTCAGCCCTGTTACCTGTTCCCAACTTTGGCCAC 1017
QY 481 CGAATTTCTCCCAATCTTAATCTTGGCTGCAAGCAAGATGCTCTCAACAGAGCTGATA 540
DB 1018 CGAATTTCTCCCAATCTTAATCTTGGCTGCAAGCAAGATGCTCTCAACAGAGCTGATA 1077
QY 541 CAGCAGAAAGGAGTTGGTATGTTAATGCCAGCTTATCCTGTCCAAAGCTGACTTT 600
DB 1078 CAGCAGAAAGGAGTTGGTATGTTAATGCCAGCTTATCCTGTCCAAAGCTGACTTT 1137
QY 601 ATCCCGAGATCTATTTCTGCGTGGCTGTAATGACAGCTTTTGTGAAGAAATTTTG 660
DB 1138 ATCCCGAGATCTATTTCTGCGTGGCTGTAATGACAGCTTTTGTGAAGAAATTTTG 1197
QY 661 CCGTGGTTGACAAATCAGTAGATTTTCAATTGAAAGCAAAAGCTCCCAATGAGATGTT 720
DB 1198 CCGTGGTTGACAAATCAGTAGATTTTCAATTGAAAGCAAAAGCTCCCAATGAGATGTT 1257
QY 721 CTAGTGAATGTTTGGTGGAGTCTCCGCTCCGCAACATGCTATGCTTACATCATG 780
DB 1258 CTAGTGAATGTTTGGTGGAGTCTCCGCTCCGCAACATGCTATGCTTACATCATG 1317
QY 781 AAGAGATGAGATGATCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGCACTACT 840

DB 1318 AAGAGATGAGATGATCTTTAGATGAAGCTTACAGATTTTGTGAAGAAAAAGCACTACT 1377
QY 841 ATATCTCCCAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAGAAG 900
DB 1378 ATATCTCCCAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAGAAG 1437
QY 901 CAGACTGAGCATCAGAGGCCAAAGCAAACTCAAGCTGCTGACCTGAGAGAGCAAT 960
DB 1438 CAGACTGAGCATCAGAGGCCAAAGCAAACTCAAGCTGCTGACCTGAGAGAGCAAT 1497
QY 961 GAACCTGTCTCTGCTGTCTTCAAGAGGTGACAGAAAGCAGACGCCCTCAGTCAACC 1020
DB 1498 GAACCTGTCTCTGCTGTCTTCAAGAGGTGACAGAAAGCAGACGCCCTCAGTCAACC 1557
QY 1021 TGTGCGCATCTGCTACCTCAGAGCAGAGGACAAAGCCCGTGCATCCCGCAGCTG 1080
DB 1558 TGTGCGCATCTGCTACCTCAGAGCAGAGGACAAAGCCCGTGCATCCCGCAGCTG 1617
QY 1081 CCGAGGTGCGCAGCGTGCAGCGCTGCTTTAGAGACAGCCCGTGTACAGGCGCTC 1140
DB 1618 CCGAGGTGCGCAGCGTGCAGCGCTGCTTTAGAGACAGCCCGTGTACAGGCGCTC 1677
QY 1141 AGTGGGCTGACCTGTCCGACAGACGCTGGAAGACCAATTAAGCTCAAGCTTCTTC 1200
DB 1678 AGTGGGCTGACCTGTCCGACAGACGCTGGAAGACCAATTAAGCTCAAGCTTCTTC 1737
QY 1201 TCTCTGATATCAATCAATGTTTCAATTCACGAGATGAGAGCATCTTACATGAGCTTC 1260
DB 1738 TCTCTGATATCAATCAATGTTTCAATTCACGAGATGAGAGCATCTTACATGAGCTTC 1797
QY 1261 TCTCTATGAAAGATGCTTTGGAATACTACAACTTTCATCTGATGAGAGCAAC 1320
DB 1798 TCTCTATGAAAGATGCTTTGGAATACTACAACTTTCATCTGATGAGAGCAAC 1857
QY 1321 AAGCTATGCTGATCTCTCCCTGTTACAGAACTATCCGAGAGATCCCGCAACAGCTCT 1380
DB 1858 AAGCTATGCTGATCTCTCCCTGTTACAGAACTATCCGAGAGATCCCGCAACAGCTCT 1917
QY 1381 GATTAAGAGAGAGCCAGCATCTCCCAAGAGCTGACACCGCAGGCTTACAGACGCA 1440
DB 1918 GATTAAGAGAGAGCCAGCATCTCCCAAGAGCTGACACCGCAGGCTTACAGACGCA 1977
QY 1441 AGCAAGGATTTGATTTGGTCAAGAACAGACAGAGTGGCAACCGCCCAAGAGTCCCTTTA 1500
DB 1978 AGCAAGGATTTGATTTGGTCAAGAACAGACAGAGTGGCAACCGCCCAAGAGTCCCTTTA 2037
QY 1501 TCTCAGCTGATGGAAGTGGAGAGCTGAGAGCAATTAACACACAGCTTCTTTTGGC 1560
DB 2038 TCTCAGCTGATGGAAGTGGAGAGCTGAGAGCAATTAACACACAGCTTCTTTTGGC 2097
QY 1561 CTTTCCACAGCCAGACAGCACTCAGAGAGTGTGCTGAGCTTGAAGGCTTGGCAC 1620
DB 2098 CTTTCCACAGCCAGACAGCACTCAGAGAGTGTGCTGAGCTTGAAGGCTTGGCAC 2157
QY 1621 TCGGATATCTTGGCCCCCAAGCTCTACCCCTTCCCTGACAGAGCTGATATTTTGGC 1680
DB 2158 TCGGATATCTTGGCCCCCAAGCTCTACCCCTTCCCTGACAGAGCTGATATTTTGGC 2217
QY 1681 ACAAGTCTCAGACTTACTGCTGAGCTGAGCATCTACGAGGAGAGTGCAGTTACT 1740
DB 2218 ACAAGTCTCAGACTTACTGCTGAGCTGAGCATCTACGAGGAGAGTGCAGTTACT 2277
QY 1741 GCTTACAGCTGACAGCCAGCTGCTTCTGCGAGACCAAGTCTTATTTGTGCGAGCGG 1800
DB 2278 GCTTACAGCTGACAGCCAGCTGCTTCTGCGAGACCAAGTCTTATTTGTGCGAGCGG 2337
QY 1801 CAGAAACCAAGTACAGAGCTGACTCGCGAGAGTGGCATGAAGAGGCCCTTTGAA 1860
DB 2338 CAGAAACCAAGTACAGAGCTGACTCGCGAGAGTGGCATGAAGAGGCCCTTTGAA 2397
QY 1861 AAGCAATTTAAACGAGAAAGCTGCAATATGGAATTTGAGAGAGCATATGACAGAAC 1920

Db	2398	AAAGCAGTTTAAACGAGAAAGCTGCCAATATGAAATTTGGAGAGAGCATATGTCAGAGAC	2457
Qy	1921	AGGTCAACGGGAAAGACCTCGGGGAAAGTTGGGACAGTCTTACGCTTTTGGGAGCATGGAA	1980
Db	2458	AGGTCAACGGGAAAGACCTCGGGGAAAGTTGGGACAGTCTTACGCTTTTGGGAGCATGGAA	2517
Qy	1981	ATCATTGAGGCTCTCC	1995
Db	2518	ATCATTGAGGCTCTCC	2532
RESULT 2			
ACC60572			
ID	ACC60572	standard; cDNA; 5450 BP.	
AC	ACC60572;		
XX			
DT	19-JUN-2003	(first entry)	
XX			
DE	Polynucleotide relating to the invention SEQ ID NO: 147.		
XX			
KW	Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic;		
KW	antiproliferative; cardiatic; cytostatic; gene therapy; liver disease;		
KW	proliferative disorder; renal failure; cardiovascular disorder;		
KW	immunological disorder; arthritis; psoriasis; congenital heart defect;		
KW	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200257460-A2.		
XX			
PD	25-JUL-2002.		
XX			
PF	20-DEC-2001; 2001WO-US050459.		
XX			
PR	20-DEC-2000; 2000US-0256868P.		
PR	30-MAR-2001; 2001US-0280186P.		
PR	01-MAY-2001; 2001US-0287735P.		
PR	05-JUN-2001; 2001US-0295848P.		
PR	25-JUN-2001; 2001US-0300465P.		
XX			
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.		
XX			
PI	Jackson DG, Feder J, Nelson T, Mintlar G, Ramenathan C, Lee L,		
PI	Siemens N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;		
PI	Kryster S, Mcatee P, Suchard S, Banaas D;		
XX			
DR	WPI; 2002-599721/64.		
XX			
PT	P-PSDB; ABR52407.		
XX			
PS	Novel polynucleotides encoding human phosphatase polypeptides useful in		
XX	the prevention or treatment of e.g. proliferative and cardiovascular		
XX	disorders.		
PS	Example 7; Fig 19; 801pp; English.		
XX			
CC	The invention relates to a novel isolated nucleic acid comprising a		
CC	polynucleotide having a nucleotide sequence selected from 40		
CC	polynucleotides fully defined in the specification. The polynucleotide of		
CC	the invention has antiproliferative, hepatotropic, nephrotropic,		
CC	antiarthritic, antiproliferative, cardiatic, and cytostatic activity. The		
CC	polynucleotide may have a use in gene therapy. A polynucleotide or		
CC	polypeptide of the invention is useful for preventing, treating or		
CC	ameliorating a medical condition, e.g. a proliferative disorder. They are		
CC	also useful for treating e.g. liver disease, renal failure, immunological		
CC	disorders including arthritis and psoriasis, cardiovascular disorders		
CC	such as congenital heart defects and congestive heart failure, and		
CC	cancer. A method of the invention is useful for diagnosing a pathological		
CC	condition or susceptibility to a condition in a subject. The present		
CC	sequence is used in the exemplification of the invention		
5Q	Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;		

Query Match	100.0%	Score 1995;	DB 6;	Length 5450;
Beet Local Similarity	100.0%	Prid. No. 0;		
Matches 1995;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGCCCATGAGATGATTGAAACTCAATTGTCTACTGAGAGGTGGTGGCTCTGCTGGAA	60	
Db	538	ATGGCCCATGAGATGATTGAACTCAATTGTCTACTGAGAGGTGGTGGCTCTGCTGGAA	597	
QY	61	AGTGAACGGAAAAAGTGTCTGCTAATTGATAGCCGGCATTTGTGGAATACATACATCC	120	
Db	598	AGTGAACGGAAAAAGTGTCTAATTGATAGCCGGCATTTGTGGAATACATACATCC	657	
QY	121	CACATTTTGGAAAGCCATTAATATCACTGCTCCAAAGCTTATGAAAGCAAGTTGCAACAG	180	
Db	658	CACATTTTGGAAAGCCATTAATATCACTGCTCCAAAGCTTATGAAAGCAAGTTGCAACAG	717	
QY	181	GACAAAGTGAATTAACAGAGCTCATCAGACATTCAGCGAAACATAAGTTGACATTGAT	240	
Db	718	GACAAAGTGAATTAACAGAGCTCATCAGACATTCAGCGAAACATAAGTTGACATTGAT	777	
QY	241	TGCAGTCAGAAAGTGTGATTTAGATCAATCAAAAGTCCCAAGATGTGTGCTCTCTCTTTCA	300	
Db	778	TGCAGTCAGAAAGTGTGATTTAGATCAAAAGTCCCAAGATGTGTGCTCTCTCTTTCA	837	
QY	301	GACTGTTTCTCACTGTATCTTGGGTTAACTGAGAAAGACTTCAACTCTGTTCACTG	360	
Db	838	GACTGTTTCTCACTGTATCTTGGGTTAACTGAGAAAGACTTCAACTCTGTTCACTG	897	
QY	361	CTTCCAGGTGGGTTTGGCTGAGTTCTCTCGTTGTTTCCCTGGCCCTGTGAAAGAAATCC	420	
Db	898	CTTCCAGGTGGGTTTGGCTGAGTTCTCTCGTTGTTTCCCTGGCCCTGTGAAAGAAATCC	957	
QY	421	ACTCTAGTCCCTACCTGCAATTTCTCAGACCTTGCTTACCTGTGGCAACATTTGGGCCAAC	480	
Db	958	ACTCTAGTCCCTACCTGCAATTTCTCAGACCTTGCTTACCTGTGGCAACATTTGGGCCAAC	1017	
QY	481	CGAATTTCTCCCAATCTTTATCTTGGCTCCAGCGAAGTGTCTCCACAAAGAGCTGATA	540	
Db	1018	CGAATTTCTCCCAATCTTTATCTTGGCTCCAGCGAAGTGTCTCCACAAAGAGAGCTGATA	1077	
QY	541	CAGAGAAATGGGAATGTGTTATGTGTAAATGCGCAGCTATACCTGTCCAAAGCCGACCTT	600	
Db	1078	CAGAGAAATGGGAATGTGTTATGTGTAAATGCGCAGCTATACCTGTCCAAAGCCGACCTT	1137	
QY	601	ATCCCCGAGTCTCATTTTCTGCGCTGTGCTGCTGTGATGACAGCTTTTGTGAAAAATTTTG	660	
Db	1138	ATCCCCGAGTCTCATTTTCTGCGCTGTGCTGTGATGACAGCTTTTGTGAAAAATTTTG	1197	
QY	661	CCGTGTGTGACAAATCAGTATGATTTTCAATGAGAAAGCAAAAGCCCTCAATGATGTGT	720	
Db	1198	CCGTGTGTGACAAATCAGTATGATTTTCAATGAGAAAGCAAAAGCCCTCAATGATGTGT	1257	
QY	721	CTAGTGCACGTGTTAGCTGGGATCTCCCCGCTCCGCCACATCGCTATTCGCTACATCATG	780	
Db	1258	CTAGTGCACGTGTTAGCTGGGATCTCCCCGCTCCGCCACATCGCTATTCGCTACATCATG	1317	
QY	781	AAGAGGATGACATGTCTTTAGATGAAAGCTTACAGATTTTGTGAAAGAAAGAAAGACTTACT	840	
Db	1318	AAGAGGATGACATGTCTTTAGATGAAAGCTTACAGATTTTGTGAAAGAAAGAAAGACTTACT	1377	
QY	841	ATATCTCCAAACTTCAATTTTCTGGGCCCAACTCTCTGAGCTATGAGAAAGATTTAGAAC	900	
Db	1378	ATATCTCCAAACTTCAATTTTCTGGGCCCAACTCTCTGAGCTATGAGAAAGATTTAGAAC	1437	
QY	901	CAGACTGAGAGATCAGAGGCCAAAGAAAGAACTCAAGCTGCTGCACTCTGAGAAAGCCAAAT	960	
Db	1438	CAGACTGAGAGATCAGAGGCCAAAGAAAGAACTCAAGCTGCTGCACTCTGAGAAAGCCAAAT	1497	
QY	961	GAACCTGTCCCTGTCTCTCAGAGAGGTGAGACAGAAAGCAGACGCCCTCAGTCCACCC	1020	
Db	1498	GAACCTGTCCCTGTCTCTCAGAGAGGTGAGACAGAAAGCAGACGCCCTCAGTCCACCC	1557	
QY	1021	TGTGCGCACTTGTCTACCTCAGAGCAGCAGGACAAAGGCCCTGTGATCCCGCCAGCGTG	1080	

```

Db      1558  TGTGCGACTCTGCTTACTCAGAGGACAGAGCAAAAGGCCCTGATCCCGCAGCGTG 1617
Qy      1081  CCCAGCGTGGCCAGCCGTCAGCCGCTGCTGTTAAGAGACACCCTGGTACAGCGCTC 1140
Db      1618  CCCAGCGTGGCCAGCCGTCAGCCGCTGCTGTTAAGAGACACCCTGGTACAGCGCTC 1677
Qy      1141  AGTGGCTGACCTGTCCGACAGACAGCTGGAAGACAGAAATAGCTCAAGGCTTC 1200
Db      1678  AGTGGCTGACCTGTCCGACAGACAGCTGGAAGACAGAAATAGCTCAAGGCTTC 1737
Qy      1201  TCTTGTGATTCAAATCAGTTTCATATTCAGCCAGCATGSCAGCATCTTACATGCTTC 1260
Db      1738  TCTGTGATATCAATCAATTCATTCATTCAGCAGATGSCAGCATCTTACATGCTTC 1797
Qy      1261  TCCATCAAGAAATGCTTTGAAATCTCAAACTTCCTCACTCTGTGATGGAACCAAC 1320
Db      1798  TCCATCAAGAAATGCTTTGAAATCTCAAACTTCCTCACTCTGTGATGGAACCAAC 1857
Qy      1321  AAGCTATGCGCACTTCCCTGTTGAGAACTATCGAGAGACGACTCCCGGAACAGTCTT 1380
Db      1858  AAGCTATGCGCACTTCCCTGTTGAGAACTATCGAGAGACGACTCCCGGAACAGTCTT 1917
Qy      1381  GATAGAGAGAGAGCCAGCATCCCAAGAACTGTCAGACCCGCGAGCTTCAGACAGCAG 1440
Db      1918  GATAGAGAGAGAGCCAGCATCCCAAGAACTGTCAGACCCGCGAGCTTCAGACAGCAG 1977
Qy      1441  AGCAAGCGATTGCTATCGGTGAGAACCGAGAGAGTGGACCGCCGAGAGGTCCCTTTTA 1500
Db      1978  AGCAAGCGATTGCTATCGGTGAGAACCGAGAGAGTGGACCGCCGAGAGGTCCCTTTTA 2037
Qy      1501  TCTCAGCTGCATCGAAGTGGAGCGGTGAGAGCAATTACCAACAGCTTCTTTTCGCG 1560
Db      2038  TCTCAGCTGCATCGAAGTGGAGCGGTGAGAGCAATTACCAACAGCTTCTTTTCGCG 2097
Qy      1561  CTTTCACACAGCCAGCAGACCTCAACAGATCTGCTGGCCCTTAAAGGCTGCGAC 1620
Db      2098  CTTTCACACAGCCAGCAGACCTCAACAGATCTGCTGGCCCTTAAAGGCTGCGAC 2157
Qy      1621  TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGAGCTGTATTTGGC 1680
Db      2158  TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGAGCTGTATTTGGC 2217
Qy      1681  ACAGAGTCTCAGACTCTTACTCTGCTGCTCAGCCATCTACGAGAGCAGTGCAGTTACTCT 1740
Db      2218  ACAGAGTCTCAGACTCTTACTCTGCTGCTCAGCCATCTACGAGAGCAGTGCAGTTACTCT 2277
Qy      1741  GCTTACAGCTGCAGCCAGCTGCCACTTGGGAGACCAAGTCTATCTGTGGCGAGCGG 1800
Db      2278  GCTTACAGCTGCAGCCAGCTGCCACTTGGGAGACCAAGTCTATCTGTGGCGAGCGG 2337
Qy      1801  CAGAGCCAGTGCAGAGCTGACTGCGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAA 1860
Db      2338  CAGAGCCAGTGCAGAGCTGACTGCGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAA 2397
Qy      1861  AAGCAGTTTAAAGCAGAAAGCTGCCAAATGGAATTTGAGAGAGACATCTTCAGAGAAC 1920
Db      2398  AAGCAGTTTAAAGCAGAAAGCTGCCAAATGGAATTTGAGAGAGACATCTTCAGAGAAC 2457
Qy      1921  AGGTCAAGGGAGAGAGCTGGGAAAAGTGGCAGTCAAGTCTTTCGGGCGAGATGAA 1980
Db      2458  AGGTCAAGGGAGAGAGCTGGGAAAAGTGGCAGTCAAGTCTTTCGGGCGAGATGAA 2517
Qy      1981  ATCATTTAGAGTCTCC 1995
Db      2518  ATCATTTAGAGTCTCC 2532

```

RESULT 3
 AAS15768
 ID AAS15768 standard; cDNA; 3059 BP.
 XX
 AC AAS15768;

```

XX      18-JUN-2002 (first entry)
DT
DE      cDNA sequence of human (dual specificity phosphatase) DUSP-10.
XX
XX      Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke;
XX      neuronal degeneration syndrome; Alzheimer's disease; depression;
XX      schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
XX      osteoporosis; diabetes; gene therapy; chromosome 12; 88.
XX
XX      Homo sapiens.
XX
XX      Key      Location/Qualifiers
XX      FT      127..2124
XX      CDS      /tag= a
XX      /product= "DUSP-10 protein"
XX
XX      WO200177340-A1.
XX
XX      18-OCT-2001.
XX
XX      06-APR-2001; 2001WO-EP003966.
XX
XX      10-APR-2000; 2000BP-00107143.
XX
XX      (MERB ) MERCK PATENT GMBH.
XX
XX      Duecker K;
XX
XX      WPI: 2002-010917/01.
XX      P-PSDB; AAU09946.
XX
XX      Novel dual specificity phosphatase polypeptides useful for treating
XX      cancer, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
XX      disease, depression, schizophrenia, asthma and immune disorders.
XX
XX      Claim 5; Page 34-37; 43pp; English.
XX
XX      The present invention relates to a new isolated dual specificity
XX      phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
XX      sequence that is fully defined in the specification. The invention also
XX      provides a sequence encoded by a 3059 nucleotide sequence fully defined
XX      in the specification, and a sequence having at least 95 % identity to the
XX      polypeptide, or fragments or variants of DUSP-10. The invention is useful
XX      for treating cancer e.g. leukaemia, colon carcinoma, lung cancer,
XX      prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
XX      stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
XX      schizophrenia, cardiac myopathies, asthma, immune disorders,
XX      inflammatory processes e.g. arthritis, bowel disease, type 1 diabetes,
XX      osteoporosis, diabetes and diabetes associated diseases. The molecules of
XX      the invention are also useful as vaccines for inducing immunological
XX      response in a mammal, in disease diagnosis and in assays for screening
XX      agonistic or antagonistic compounds. Other uses of the invention include
XX      identifying membrane bound or soluble receptors, as a diagnostic reagent,
XX      in chromosome localisation studies, and as a valuable tool in tissue
XX      expression studies. The present sequence represents cDNA of the human
XX      dual specificity phosphatase, DUSP-10. DUSP-10 is located on chromosome
XX      12

```

SO Sequence 3059 BP; 831 A; 761 C; 709 G; 758 T; 0 U; 0 Other;

Query Match 94.9%; Score 1893; DB 6; Length 3059;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy      1  ATGGCCCATGAGATGATGGAATCTCAATTTGTACTGAGAGTGTGGCTCTGCTGGAA 60
Db      127  ATGGCCCATGAGATGATGGAATCTCAATTTGTACTGAGAGTGTGGCTCTGCTGGAA 186
Qy      61  AGTGAACGGAAGAAAGTCTGCTGTAATGATCCGCGCATTTGTGGAAATACATACATCC 120
Db      187  AGTGAACGGAAGAAAGTCTGCTGTAATGATCCGCGCATTTGTGGAAATACATACATCC 246

```

QY	121	CACATTTTGGAGCCATTAAATATCAACTGCTCCAAAGCTTAATGAAGCCAAAGTTGCACAG	180
Db	247	CACATTTTGGAGCCATTAAATATCAACTGCTCCAAAGCTTAATGAAGCCAAAGTTGCACAG	306
QY	181	GACAAAGTGTAAATPACAGAGCTCATCCAGCAATTCAGGAAACATPAAAGTTGACATTGAT	240
Db	307	GACAAAGTGTAAATPACAGAGCTCATCCAGCAATTCAGGAAACATPAAAGTTGACATTGAT	366
QY	241	TGCAGTCAGAAAGTGTAGTTTACGATCAAAAGCTCCAAAGTGTGCTCTCTCTTCA	300
Db	367	TGCAGTCAGAAAGTGTAGTTTACGATCAAAAGCTCCAAAGTGTGCTCTCTCTTCA	426
QY	301	GACTGTTTTCTCAGCTACTCTTGGGTAACTCGAAGAAAGCTTCACTGTTCACCTG	360
Db	427	GACTGTTTTCTCAGCTACTCTTGGGTAACTCGAAGAAAGCTTCACTGTTCACCTG	486
QY	361	CTTGCAAGGTGGGTGGTGGTGAATCTCTGCTGTTTCCCTGGCCCTGTGGAAGAAATCC	420
Db	487	CTTGCAAGGTGGGTGGTGGTGAATCTCTGCTGTTTCCCTGGCCCTGTGGAAGAAATCC	546
QY	421	ACTCTAGTCCCTACCTGCATTTCTCAGCCCTGCTTACCTGTGTCCCAATTTGGCCAAAC	480
Db	547	ACTCTAGTCCCTACCTGCATTTCTCAGCCCTGCTTACCTGTGTGTCCCAATTTGGCCAAAC	606
QY	481	GCAATTTCTTCCCAATCTTTATCTTGCGCTGCCAGCGAATGTCTCACAAGAGCTGATA	540
Db	607	GCAATTTCTTCCCAATCTTTATCTTGCGCTGCCAGCGAATGTCTCACAAGAGCTGATG	666
QY	541	CAGCAGAAATGGATGGTATGTGTGTTAATGCAAGCTPACTGTGTCGCAAGCCGACCTT	600
Db	667	CAGCAGAAATGGATGGTATGTGTGTTAATGCAAGCTPACTGTGTCGCAAGCCGACCTT	726
QY	601	ATCCCCGAGTCTCATTTCTCGCGTGTGCTGTGATGACAGCTTTTGTGAAGAAATTTTG	660
Db	727	ATCCCCGAGTCTCATTTCTCGCGTGTGCTGTGATGACAGCTTTTGTGAAGAAATTTTG	786
QY	661	CCGTGTTTGGACAATACATGATATTCAATTGAAGAAAGCAAAAGCTCCATAGATGTGTT	720
Db	787	CCGTGTTTGGACAATACATGATATTCAATTGAAGAAAGCAAAAGCTCCATAGATGTGTT	846
QY	721	CTAGTGCACGTGTTTACGTGGGATCTCCGCGTCCGCACTCAATGCGCTACATCATG	780
Db	847	CTAGTGCACGTGTTTACGTGGGATCTCCGCGTCCGCACTCAATGCGCTACATCATG	906
QY	781	AAGAGATGACATGCTCTTTAGATGAACTTACAGATTTGTGAAGAAAGAAAGAAAGCTACT	840
Db	907	AAGAGATGACATGCTCTTTAGATGAACTTACAGATTTGTGTGAAGAAAGAAAGAAAGCTACT	966
QY	841	ATATCTCCAAACTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTPAAGAC	900
Db	967	ATATCTCCAAACTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTPAAGAC	1026
QY	901	CAGACTGAGATCAAGGGCCAAAGGCAAACTCAAGCTGTGACACTGTGGAAGAACCAAT	960
Db	1027	CAGACTGAGATCAAGGGCCAAAGGCAAACTCAAGCTGTGACACTGTGGAAGAACCAAT	1086
QY	961	GAACTGTCCCTGCTGTCTCAGAGGGGTGAGACAGAAACCGAAGCGCCCTCAGTCCACC	1020
Db	1087	GAACTGTCCCTGCTGTCTCAGAGGGGTGAGACAGAAACCGAAGCGCCCTCAGTCCACC	1146
QY	1021	TGTGCGCACTGTCTACCTCAGAGGCAAGGCAAAAGCCCGTGCATCCCGCCAGCGTG	1080
Db	1147	TGTGCGCACTGTCTACCTCAGAGGCAAGGCAAAAGCCCGTGCATCCCGCCAGCGTG	1206
QY	1081	CCCAACGTGCCAGGCTGCAAGCCGTGCTGTTAAGGACAGCCCGCTGTGTCAAGCGCTC	1140
Db	1207	CCCAACGTGCCAGGCTGCAAGCCGTGCTGTTAAGGACAGCCCGCTGTGTCAAGCGCTC	1266
QY	1141	AGTGGGCTGCACCTGTCCGACAGACAGGCTGGAAGACACATPAAAGCTCAACGTTCCCTC	1200
Db	1267	AGTGGGCTGCACCTGTCCGACAGACAGGCTGGAAGACACATPAAAGCTCAACGTTCCCTC	1326
QY	1201	TCTGTGATATCAAAATCAGTTTCATATTTCAAGCAGATGCGACATCTCTTACATGCTTTC	1260

Dd	1327	TCTCTGGATATCAATACAGTTTCTCATTTACAGCCAGATGGCAGCATCTTACATGAGCTTC	1386
Qy	1261	TCCCTCATCAGAAAGTCTTTGGAAATCTAAGAACTTTCACATCTCTGATGGGACCAAC	1320
Dd	1387	TCCCTCATCAGAAAGATGGCTTTGGAAATCTAAGAACTTTCACATCTCTGATGGGACCAAC	1446
Qy	1321	AAGCTATGCCAGTTCTCTCCCTGTTCAGAACTATCGAGCAGACTCCGAAACCAAGTCT	1380
Dd	1447	AAGCTATGCCAGTTCTCTCCCTGTTCAGAACTATCGAGCAGACTCCGAAACCAAGTCT	1506
Qy	1381	GATTAAGAGAGAGCAGACATCCCAAGAAAGCTGAGCGCCAGGCGCTTCAGACAGCAG	1440
Dd	1507	GATTAAGAGAGAGCAGACATCCCAAGAAAGCTGAGCGCCAGGCGCTTCAGAGAGCAG	1566
Qy	1441	AGCAAGGGAATTTGACTTTGGTCTAGAAACAGACAGAGTGGACCGCCAGAGGTCCCTTTTA	1500
Dd	1567	AGCAAGGGAATTTGACTTTGGTCTAGAAACAGACAGAGTGGACCGCCAGAGGTCCCTTTTA	1626
Qy	1501	TCTCCACTGCATCGAAGTGGAGGCTGGAGAGCAATTAACAACACACAGCTTCTTTTGGC	1560
Dd	1627	TCTCCACTGCATCGAAGTGGAGGCTGGAGAGCAATTAACAACACACAGCTTCTTTTGGC	1686
Qy	1561	CTTTTCCACACGCGCAGACACCTTCACGAAGCTGCTGGCCCTGGGCGCTTAAGGGCTGGCAC	1620
Dd	1687	CTTTTCCACACGCGCAGACACCTTCACGAAGCTGCTGGCCCTGGGCGCTTAAGGGCTGGCAC	1746
Qy	1621	TCGGATATCTTTGGGCCCCCAGACCTCTACCCCTTCCCTGACACAGACGCTGTAATTTTGGC	1680
Dd	1747	TCGGATATCTTTGGGCCCCCAGACCTCTACCCCTTCCCTGACACAGACGCTGTAATTTTGGC	1806
Qy	1681	ACAGAGTCTCTCACACTTCTACTCTGCTCAGCCATCTACGAGAGGAGTGCACGTTACTCT	1740
Dd	1807	ACAGAGTCTCTCACACTTCTACTCTGCTCAGCCATCTACGAGAGGAGTGCACGTTACTCT	1866
Qy	1741	GCCATACAGCTGCAGCCAGCTGGCCCACTTTGGGAGAGCCAACTTATTTCTGGGCCAGGCGG	1800
Dd	1867	GCCATACAGCTGCAGCCAGCTGGCCCACTTTGGGAGAGCCAACTTATTTCTGGGCCAGGCGG	1926
Qy	1801	CAGAAAGCCAAAGTCAGAGCTGACTCGCGCGGAGCTGGGATGAAAGAGAGAGCCCTTTGAA	1860
Dd	1927	CAGAAAGCCAAAGTCAGAGCTGACTCGCGCGGAGCTGGGATGAAAGAGAGAGCCCTTTGAA	1986
Qy	1861	AAGCAGTTTAAACGACAGAGCTGCAATTTGAGAGAGCATCATGTACAGAGAAC	1920
Dd	1987	AAGCAGTTTAAACGACAGAGCTGCAATTTGAGAGAGCATCATGTACAGAGAAC	2046
Qy	1921	AGGTTCACGGGAAAGCTGGGGAAAGTGGGCAATGCTTCTTTGGGCGACGATGGAA	1980
Dd	2047	AGGTTCACGGGAAAGCTGGGGAAAGTGGGCAATGCTTCTTTGGGCGACGATGGAA	2106
Qy	1981	ATCATTTAGGCTCTCC	1995
Dd	2107	ATCATTTAGGCTCTCC	2121
RESULT 4			
ABK47596			
ID	ABK47596 standard; cDNA; 3496 BP.		
XX	ABK47596;		
AC			
XX			
DT	02-JUL-2002 (first entry)		
XX			
DE	cDNA encoding human dual-specificity phosphatase-3 (DSP-16) protein.		
XX			
KM	Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;		
KM	mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;		
KM	cancer; graft-versus-host disease; allergy; metabolic disease;		
KM	abnormal cell growth; abnormal cell proliferation; contact inhibition;		
KM	cell cycle abnormality; anchorage independent cell growth; apoptosis;		
KM	intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; 86.		
XX			

OS	Homo sapiens.
FH	Key
FX	Location/Qualifiers
FT	CDS
FT	562..2559
FT	/tag= a
FT	/product= "Human dual-specificity phosphatase-3 (DSP-16)
FT	protein"
XX	
XX	MO200226997-A2.
XX	
XX	04-APR-2002.
PD	
PF	25-SEP-2001; 2001MO-US031024.
XX	
XX	26-SEP-2000; 2000US-0235487P.
FR	(CEPT-) CEPTYR INC.
PA	
P1	Luche RM, Wei B;
XX	
XX	WPI; 2002-315802/35.
DR	P-PsDB; AAU79156.
DR	
XX	New DSP-16 polypeptide, useful for identifying modulators of its
PT	activity, which can be used in the treatment of disorders such as
XX	Duchenne muscular dystrophy, or cancer.
PS	Claim 7; Fig 1; 87pp; English.
XX	
CC	The present invention relates to a new polypeptide, DSP-16, having a 665
CC	amino acid sequence, given in the specification, or a variant having at
CC	least 50 % identical residues, which retains the ability to
CC	dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
CC	invention can be used for identifying agents which modulate DSP-16
CC	activity, for modulation of a proliferative response in a cell, survival
CC	of a cell, or differentiation of a cell. The cell displays contact
CC	inhibition of cell growth or anchorage independent growth and may display
CC	altered intercellular adhesion. The agent may modulate apoptosis, or the
CC	cell cycle. The identified modulators can be used to treat Duchenne
CC	muscular dystrophy, cancer, graft-versus-host disease, autoimmune
CC	diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
CC	cell proliferation, and cell cycle abnormalities. The present nucleic
CC	acid sequence is that of the human dual-specificity phosphatase-3 (DSP-
CC	16) gene located on chromosome 12p. This sequence encodes the human DSP-
CC	16 protein of the invention
XX	
XX	
SQ	Sequence 3496 BP; 939 A; 854 C; 856 G; 847 T; 0 U; 0 Other;
Query Match	94.9%; Score 1893; DB 6; Length 3496;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 1993; Conservative	0; Mismatches 2; Indels 0; Gaps 0
OY	1 ATGGCCCATGATGATGTGAACCTCAATTGTTACTGAGAGTGTGGCTGCTGTGAA 60
Db	562 ATGGCCCATGATGATGTGAACCTCAAAVTGTTACTGAGAAGTGTTGGCTGCTGGA 621
OY	61 AGTGAACCGAAAAAAGTCGCTAATTATGATGACCGGCATTTGGTAATCAATACATCC 120
Db	622 AGTGAACCGAAAAAAGTCGCTAATTATGATGACCGGCATTTGGTAATCAATACATCC 681
OY	121 CA CATTTTGGAACCATTAATAATCAACTGCTCAAAGCTTATGAAGCAAGTTGCCAAGC 180
Db	682 CACATTTTGGAAAGCCATTAATATCAACTGCTCAAAGCTTATGAAGCAAGTTGCCAAGC 741
OY	181 GACAAGSTTTAATTATACAGAGCTCATCCAGATTCACGCGAAACATAAGTTGACATTGAT 240
Db	742 GACAAGSTTTAATTATACAGAGCTCATCCAGATTCACGCGAAACATAAGTTGACATTGAT 801
OY	241 TGCAAGTCAGAAAGTTGTAGTTTAGATCAATCAAGCTCCCAAGATGTCCTCTCTCTTCA 300
Db	802 TGCAAGTCAGAAAGTTGTAGTTTAGATCAATCAAGCTCCCAAGATGTCCTCTCTCTTCA 861
OY	301 GACCTGTTTTCTCACTGTACTTCTGGGTTAAACTGGAGAAAGAGCTTCAACTGTGTCACTG 360

Db	862	GACTGTTTTCTCACTGTATCTCTGGGTTAACTGGAGAAAGACTTCAACTCTGTTCACTG	921
OY	361	CTTGACAGTGGGGTTCGCTGAGATTCTCGATGGTTTCCCTGGCCCTCTGTGAAGAAATCC	420
Db	922	CTTGACAGTGGGGTTCGCTGAGATTCTCGATGGTTTCCCTGGCCCTCTGTGAAGAAATCC	981
OY	421	ACTCTAGTCCCTTACCTGCATTTTCTCAGCCTTGCTTACCTGTTCGCAACATTTGGGCCAAC	480
Db	982	ACTCTAGTCCCTTACCTGCATTTTCTCAGCCTTGCTTACCTGTTCGCAACATTTGGGCCAAC	1041
OY	481	CGAATTTCTTCCCAATCTTTATCTTGGGTGCAACGAGATGTCCTCAACAAGAGCTGAT	540
Db	1042	CGAATTTCTTCCCAATCTTTATCTTGGGTGCAACGAGATGTCCTCAACAAGAGCTGATG	1101
OY	541	CAGCAGAAATGGGATGGTATATGTTTAAATGCCAGTATACCTGTGCAAGCCTGACTTT	600
Db	1102	CAGCAGAAATGGGATGGTATATGTTTAAATGCCAGTATACCTGTGCAAGCCTGACTTT	1161
OY	601	ATCCCGAGCTCATTTCCCTGGGTGCTCGTGAATGACGACTTTTGTGAGAAAATTTTG	660
Db	1162	ATCCCGAGCTCATTTCCCTGGGTGCTCGTGAATGACGACTTTTGTGAGAAAATTTTG	1221
OY	661	CCGTGGTTGGACAAATCAGTAAATTTTATTGAGAAAGCAAAGCCTCCAAATGGATGTGT	720
Db	1222	CCGTGGTTGGACAAATCAGTAAATTTTATTGAGAAAGCAAAGCCTCCAAATGGATGTGT	1281
OY	721	CTATGTCACCTGTTTATGCTGGGATCTCCGCTCCGCCACCATGCTATGCTTACATCATG	780
Db	1282	CTATGTCACCTGTTTATGCTGGGATCTCCGCTCCGCCACCATGCTATGCTTACATCATG	1341
OY	781	AAGAGATGAGATGCTTTAGATGAATGAACCTTACAGATTTGTGAAAGAAAAGAACTACT	840
Db	1342	AAGAGATGAGATGCTTTAGATGAATGAACCTTACAGATTTGTGAAAGAAAAGAACTACT	1401
OY	841	ATATCTCCCAACTTTCAATTTTCTGGGCCAACTCTTGATGTAAGAAAGAAATTTAAGAC	900
Db	1402	ATATCTCCCAACTTTCAATTTTCTGGGCCAACTCTTGATGTAAGAAAGAAATTTAAGAC	1461
OY	901	CAGACTGAGCATTCAGGGCCCAAAGAGCAAATCTCAAGCTGCTGCACCTGAGAAACCTAAT	960
Db	1462	CAGACTGAGCATTCAGGGCCCAAAGAGCAAATCTCAAGCTGCTGCACCTGAGAAACCTAAT	1521
OY	961	GAACTGTTCCTGCTGTCTCAAGAGGTGACAGAAAACGAGACGCCCTCAGTCCACCC	1020
Db	1522	GAACTGTTCCTGCTGTCTCAAGAGGTGACAGAAAACGAGACGCCCTCAGTCCACCC	1581
OY	1021	TGTGCGCACTCTGCTACTCAGAGGACGACGAAAGAGCCCGCTGATCCCCGACGCTG	1080
Db	1582	TGTGCGCACTCTGCTACTCAGAGGACGACGAAAGAGCCCGCTGATCCCCGACGCTG	1641
OY	1081	CCGACGCTGCCAGCGTGCAGCCGCTGTTAAGAGCAGCCCGCTGTATCAGAGGCTC	1140
Db	1642	CCGACGCTGCCAGCGTGCAGCCGCTGTTAAGAGCAGCCCGCTGTATCAGAGGCTC	1701
OY	1141	AGTGGGCTGCACCTGTGCCGAGACAGGCTTGGAAAGACGCAATAGCTCAAGGCTTCTTC	1200
Db	1702	AGTGGGCTGCACCTGTGCCGAGACAGGCTTGGAAAGACGCAATAGCTCAAGGCTTCTTC	1761
OY	1201	TCTCTGGATATCAAAATCAGTTTCAATTTACGCCAGATGAGCAGATCTTTACATGGCTTC	1260
Db	1762	TCTCTGGATATCAAAATCAGTTTCAATTTACGCCAGATGAGCAGATCTTTACATGGCTTC	1821
OY	1261	TTCCTCATCAAAATGTTTGGAAATCTACAAACCTTTCACATATCTGTGATGGGACCAAC	1320
Db	1822	TTCCTCATCAAAATGTTTGGAAATCTACAAACCTTTCACATATCTGTGATGGGACCAAC	1881
OY	1321	AAGCTATGCCAGTTTCCCTGTTCAGAAACTATGGAGCACAATCCGAAACAGTCTCT	1380
Db	1882	AAGCTATGCCAGTTTCCCTGTTCAGAAACTATGGAGCACAATCCGAAACAGTCTCT	1941
OY	1381	GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGCAGACCGCCAGGCTTTACAGACGCCAG	1440

QY 481 GGAATTCCTCCCAATCTTTATCTTGTGCTGCGACGAGATGCTCTCAACAGAGAGCTGATA 540
 DB 1069 CCAATTCCTCCCAATCTTTATCTTGTGCTGCGACGAGATGCTCTCAACAGAGAGCTGATA 1128
 QY 541 CAGCGAAGTGGATGTTATGTTAATGCGACCTATACCTGTCCAAAGCTGACTTT 600
 DB 1129 CAGCGAAGTGGATGTTATGTTAATGCGACCAATACCTGTCCAAAGCTGACTTT 1188
 QY 601 ATCCCGAGTCTCATTTCCGCTGCTGCTGGAATGACAGCTTTTGGAAAAATTTTG 660
 DB 1189 ATCCCGAGTCTCATTTCCGCTGCTGCTGGAATGACAGCTTTTGGAAAAATTTTG 1248
 QY 661 CCGTGGTTGGACAATTCAGATATTTCTTGAAGAAAGCAAAAGCTCCATATGATGTGT 720
 DB 1249 CCGTGGTTGGACAATTCAGATATTTCTTGAAGAAAGCAAAAGCTCCATATGATGTGT 1308
 QY 721 CTAGTGCACGTGTTAGCTGGGATCTCCCGCTCCGACATCGCTATCGCTACATCATG 780
 DB 1309 CTAGTGCACGTGTTAGCTGGGATCTCCCGCTCCGACATCGCTATCGCTACATCATG 1368
 QY 781 AAGAGATGACATGCTCTTATGATGAAGCTTAACAGATTTGTGAAGAAAAAGCCTACT 840
 DB 1369 AAGAGATGACATGCTCTTATGATGAAGCTTAACAGATTTGTGAAGAAAAAGCCTACT 1428
 QY 841 ATATCTCCAACTTCAATTTTCTGCGCAACTCTCTGACTATGAGAAAGATTTAAGAC 900
 DB 1429 ATATCTCCAACTTCAATTTTCTGCGCAACTCTCTGACTATGAGAAAGATTTAAGAC 1488
 QY 901 CAGACTGAGCATCAGGGCCCAAGAGCAAACTCAAGTGTGCACTCGGAGAAAGCCCAAT 960
 DB 1489 CAGACTGAGCATCAGGGCCCAAGAGCAAACTCAAGTGTGCACTCGGAGAAAGCCCAAT 1548
 QY 961 GAACCTGCTCCCTGCTGCTCTCAAGAGGTGACAGAAAAAGCAGACGCCCTCAGTCCACC 1020
 DB 1549 GAACCTGCTCCCTGCTGCTCTCAAGAGGTGACAGAAAAAGCAGACGCCCTCAGTCCACC 1608
 QY 1021 TGTGCGACTCTGCTACTCTCAGAGGCGACGACGACAAAGGCCGCTGCACTCCGACGGTG 1080
 DB 1609 TGTGCGACTCTGCTACTCTCAGAGGCGACGACGACAAAGGCCGCTGCACTCCGACGGTG 1668
 QY 1081 CCCAGCGTGGCCAGCGCTGACAGCGCTGCTGTAGAGAGCAGCCCGGTGGATCAGGGCGTC 1140
 DB 1669 CCCAGCGTGGCCAGCGCTGACAGCGCTGCTGTAGAGAGCAGCCCGGTGGATCAGGGCGTC 1728
 QY 1141 AGTGGGTGCACTGTGTCGACAGACAGGCTGGAGAGACAGCAATTAAGCTCAAGCGTTCTTC 1200
 DB 1729 AGTGGGTGCACTGTGTCGACAGACAGGCTGGAGAGACAGCAATTAAGCTCAAGCGTTCTTC 1788
 QY 1201 TCTCTGATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1260
 DB 1789 TCTCTGATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1848
 QY 1261 TCTCTCATCAGAGATGCTTTGAAATACTACAAACCTTCCACTACTGTGATGGAGCAAC 1320
 DB 1849 TCTCTCATCAGAGATGCTTTGAAATACTACAAACCTTCCACTACTGTGATGGAGCAAC 1908
 QY 1321 AAGCTATGCCAGTTTCCCTCTGTTCAAGAACTATCGGAGCAACTCCGAAACCAAGTCT 1380
 DB 1909 AAGCTATGCCAGTTTCCCTCTGTTCAAGAACTATCGGAGCAACTCCGAAACCAAGTCT 1968
 QY 1381 GATTAAGAGAGAGCCAGCATCCCAAGAGCTGACAGACCGGCAAGGCTTCAACAGCCAG 1440
 DB 1969 GATTAAGAGAGAGCCAGCATCCCAAGAGCTGACAGACCGGCAAGGCTTCAACAGCCAG 2028
 QY 1441 AGCAAGCATTTGATTCGCTCAGAACCAAGCAGAGTGGACCGCCAGAGGTCCCTTTTA 1500
 DB 2029 AGCAAGCATTTGATTCGCTCAGAACCAAGCAGAGTGGACCGCCAGAGGTCCCTTTTA 2088
 QY 1501 TCTCCACTGATCGAAGTGGAGCGTGGAGGACATTTACACACAGCTTCTTTTGGGC 1560
 DB 2089 TCTCCACTGATCGAAGTGGAGCGTGGAGGACATTTACACACAGCTTCTTTTGGGC 2148
 QY 1561 CTTTCACCAAGCAGACAGCACTCAGAAAGTCTGAGCTTGAAGGCGTGGAC 1620

DB 2149 CTTTCAACAGCCAGAGACAGCACTCAGAAAGTCTGCTGGCCTGGCCTTAAGGGCTGGCAC 2208
 QY 1621 TCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGAGCTGTAATTTGCC 1680
 DB 2209 TCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGAGCTGTAATTTGCC 2268
 QY 1681 ACAGAGTCTCACACTTTACTCTGCTGCTCAGCATTCTACGAGGAGAGTCCAGTTACTCT 1740
 DB 2269 ACAGAGTCTCACACTTTACTCTGCTGCTCAGCATTCTACGAGGAGAGTCCAGTTACTCT 2328
 QY 1741 GCCTACAGCTGACAGCAGCTGCGCACTTGCGAGACCAAGTCTATCTGTGCGGAGCGG 1800
 DB 2329 GCCTACAGCTGACAGCAGCTGCGCACTTGCGAGACCAAGTCTATCTGTGCGGAGCGG 2388
 QY 1801 CAGAACCAAGTACAGAGCTGACTCGCGGCGAGAGCTGGCATGAGAGAGCCCTTTGAA 1860
 DB 2389 CAGAACCAAGTACAGAGCTGACTCGCGGCGAGAGCTGGCATGAGAGAGCCCTTTGAA 2448
 QY 1861 AAGCATTTTAAACGACAAAGCTGCCAATTTGAAATTTGAGAGAGCATCATGTCAAGAAC 1920
 DB 2449 AAGCATTTTAAACGACAAAGCTGCCAATTTGAAATTTGAGAGAGCATCATGTCAAGAAC 2508
 QY 1921 AGGTACGAGAAAGAGCTGAGGAAAGTGGGACAGTCACTAGCTTTTGCGGAGAGTGA 1980
 DB 2509 AGGTACGAGAAAGAGCTGAGGAAAGTGGGACAGTCACTAGCTTTTGCGGAGAGTGA 2568
 QY 1981 ATCATTGAGTCTCC 1995
 DB 2569 ATCATTGAGTCTCC 2583

RESULT 6
 ABR49402
 ID ABR49402 standard; cDNA; 3544 BP.
 XX ABR49402;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE cDNA encoding human dual specificity phosphatase 21117 protein.
 XX
 KW Human; dual specificity phosphatase 21117; erythroid-related disorder;
 KW haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
 KW erythrocytosis; liver-related disorder; cancer; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 589..2586
 FT /tag= a
 FT /product= "Human dual specificity phosphatase 21117"
 FT /note= "Specifically claimed in claim 2"
 XX
 EN US2002034807-A1.
 XX
 PD 21-MAR-2002.
 XX
 PP 23-MAR-2001; 2001US-00816494.
 XX
 PR 24-MAR-2000; 2000US-0191858P.
 XX
 PA (MEYE/) MEYERS R. A.
 XX
 PI Meyers RA;
 XX
 DR WPI; 2002-351088/38.
 DR P-PSDB; AAU79929.
 XX
 XX New nucleic acids, designated 38692 and 21117, encoding dual specificity
 PT phosphatases for treating cell proliferation and differentiation
 PT disorders including hematopoietic and erythroid-related disorders and
 PT cancers.

DB 2449 AAGCAGTTTAAACGACGAACTGCCAATGAAATTGGAGAGACATCATGTCAGAGAAC 2508
QY 1921 AGCTCAGCGGAAAGAGCTGGGAAAGTGGGCACTCACTGTTTCCGCGACGATGGAA 1980
DB 2509 AGCTCAGCGGAAAGAGCTGGGAAAGTGGGCACTCACTGTTTCCGCGACGATGGAA 2568
QY 1981 ATCATTGAGGTCTCC 1995
DB 2569 ATCATTGAGGTCTCC 2583
RESULT 7
ABK14474
ID ABK14474 standard; cDNA; 3766 BP.
XX
AC ABK14474;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human protein phosphatase 7 (PP7) cDNA sequence.
XX
KW Human; Protein phosphatase 7; PP7; immune system disorder; AIDS;
KW acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
KW Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
KW dementia; Parkinson's disease; developmental disorder; Down's syndrome;
KW cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
KW melanoma; myeloma sarcoma; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 538..2535
FT CDS /*tag= a
FT /product= "Protein_phosphatase_7_(PP7)"
XX
XX WO200210363-A2.
XX
XX 07-FEB-2002.
XX
XX 26-JUL-2001; 2001WO-US023716.
XX
XX 28-JUL-2000; 2000US-0221679P.
XX 03-AUG-2000; 2000US-0223272P.
XX 10-AUG-2000; 2000US-0224309P.
XX 18-AUG-2000; 2000US-0226728P.
XX 30-AUG-2000; 2000US-0229254P.
XX 08-SEP-2000; 2000US-0231366P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;
PI Stewart EA, Gandhi AR, Patterson C, Lee EA, Hafalia AUA, Lu DM;
PI Tribouley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB,
PI Walla NK, Kearney L;
XX
DR WPI: 2002-188735/24.
DR P-PSDB; AAU75789.
XX
XX
PT New protein phosphatases, useful for diagnosing, treating or preventing
PT immune system disorders (e.g. Crohn's disease), neurological disorders
PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
PT cancers).
XX
XX Claim 5; Page 114-115, 117pp; English.
XX
XX The present invention relates to a new polypeptide, a naturally occurring
XX amino acid sequence at least 95 % identical to it, a biologically active
XX fragment of it or an immunogenic fragment of it. The polypeptides,
XX polynucleotides, agonists and antagonists are useful for diagnosing,
XX treating or preventing disorders associated with aberrant expression of
XX protein phosphatases (PP), particularly immune system disorders e.g.
XX acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
XX asthma or Crohn's disease, neurological disorders e.g. epilepsy.

CC Huntington's disease, dementia or Parkinson's disease, developmental
CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
CC or sarcoma. The present nucleic acid sequence encodes human protein
CC phosphatase 7 (PP7) which is one of several human protein phosphatases
CC (AAU75783-AAU75792) of the invention
XX
SQ Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;
Query Match 94.9%; Score 1893; DB 6; Length 3766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATGTAAGAACTCAAAATTGTAAGAGAGGTGGCTGTGGAA 60
DB 538 ATGGCCCATGAGATGATGTAAGAACTCAAAATTGTAAGAGAGGTGGCTGTGGAA 597
QY 61 AGTGAACGGAAAAAGTGTCTAATTGATAGCCGGCCATTGTGTGAAATACATATCC 120
DB 598 AGTGAACGGAAAAAGTGTCTAATTGATAGCCGGCCATTGTGTGAAATACATATCC 657
QY 121 CACATTTTGAAGCCATTAAATCACTGCTTCAAGTTTGAAGCGAAGTTGCAACAG 180
DB 658 CACATTTTGAAGCCATTAAATCACTGCTTCAAGTTTGAAGCGAAGTTGCAACAG 717
QY 181 GACAAAGTTTAAATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGAATTGAT 240
DB 718 GACAAAGTTTAAATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGTTGAATTGAT 777
QY 241 TGCAGTCAGAAAGTTGTAGTTTACATCAAGAGCTCCCAAGATGTCCTCTCTTCA 300
DB 778 TGCAGTCAGAAAGTTGTAGTTTACATCAAGAGCTCCCAAGATGTCCTCTCTTCA 837
QY 301 GACTGTTTCTCAGTGAATCTTGGGTAACGTGAGAAAGAGCTTCACTCTTTCACCTG 360
DB 838 GACTGTTTCTCAGTGAATCTTGGGTAACGTGAGAAAGAGCTTCACTCTTTCACCTG 897
QY 361 CTTCGAGGTGGGTTTGCAGATTTCTCTGTTGTTTCCCTGGCCCTCTGTGAAGAAATCC 420
DB 898 CTTCGAGGTGGGTTTGCAGATTTCTCTGTTGTTTCCCTGGCCCTCTGTGAAGAAATCC 957
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGTCTTACCTGTGCGCAATTGGGCGAAC 480
DB 958 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGTCTTACCTGTGCGCAATTGGGCGAAC 1017
QY 481 CGAATTTCTCCCAATCTTATCTTGGCTGCCAGGAGATGCTTCAACAAAGAGCTGATA 540
DB 1018 CGAATTTCTCCCAATCTTATCTTGGCTGCCAGGAGATGCTTCAACAAAGAGCTGATA 1077
QY 541 CAGCAGAAATGGGATTTGTTAATGTTAAATGCCAGCTAATCCGTCCAAAGCCTGACTTT 600
DB 1078 CAGCAGAAATGGGATTTGTTAATGTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTT 1137
QY 601 ATCCCGAGATCTCAATTTCTGCGTGTGCTGTGAATACAGACTTTTGTGAGAAATTTTG 660
DB 1138 ATCCCGAGATCTCAATTTCTGCGTGTGCTGTGAATACAGACTTTTGTGAGAAATTTTG 1197
QY 661 CCGTGTGTTGACAAATAGTATGATTTTATGAGAAAGCAAAAGCCTTCAATGATGTGTT 720
DB 1198 CCGTGTGTTGACAAATAGTATGATTTTATGAGAAAGCAAAAGCCTTCAATGATGTGTT 1257
QY 721 CTAGTGCACTGTTTAAAGTGGGATCTCCGCTCGCGCAACCATGCTATGCTTATCATCATG 780
DB 1258 CTAGTGCACTGTTTAAAGTGGGATCTCCGCTCGCGCAACCATGCTTATGCTTATCATCATG 1317
QY 781 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTTACT 840
DB 1318 AAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTTACT 1377
QY 841 ATATCTTCCAAACTTCAATTTTCTGGGCAACTCTGGAATGAGAAAGATTAAGAAC 900
DB 1378 ATATCTTCCAAACTTCAATTTTCTGGGCAACTCTGGAATGAGAAAGATTAAGAAC 1437


```
Db 364 GACAAAGTGTAAATTACAGAGCTCATCCAGGATTCAGGGAACATTAAGTTGACATTGAT 423
Qy 241 TGCAGTCAGAAAGTGTGATTTAGATCAATCAAGCTCCCAAGTGTTCCTCTCTCTTCA 300
Db 424 TGCAGTCAGAAAGTGTGATTTAGATCAATCAAGCTCCCAAGTGTTCCTCTCTCTTCA 483
Qy 301 GACTGTTTCTCACTGTAATCTCTGAGTAACTGAGAAAGAGCTCAACTCTGTTCACTG 360
Db 484 GACTGTTTCTCACTGTAATCTCTGAGTAACTGAGAAAGAGCTCAACTCTGTTCACTG 543
Qy 361 CTTCGAGGTGGGTTTGTGAGTCTCTGTTGTTTCCCTGAGCTCTGAGAAAGAAATCC 420
Db 544 CTTCGAGGTGGGTTTGTGAGTCTCTGTTGTTTCCCTGAGCTCTGAGAAAGAAATCC 603
Qy 421 ACTGAGTCCCACTGATTCAGCTTCTCAGCTTCTTCACTGTTGCAACATTTGGCCAAC 480
Db 604 ACTGAGTCCCACTGATTCAGCTTCTCAGCTTCTTCACTGTTGCAACATTTGGCCAAC 663
Qy 481 CGAATTCCTCCCAATCTTTATCTTGAGCTGCAGGAGATGCTCTCAACAGAGAGTGATA 540
Db 664 CGAATTCCTCCCAATCTTTATCTTGAGCTGCAGGAGATGCTCTCAACAGAGAGTGATA 723
Qy 541 CAGCAGAAATGGGATTTGGTTATGTGTAATGCAAGTAACTGTGCCAAAGCCTGACTTT 600
Db 724 CAGCAGAAATGGGATTTGGTTATGTGTAATGCAAGTAACTGTGCCAAAGCCTGACTTT 783
Qy 601 ATCCCGAGTCTCATTTCCGCGCTGCTGAGTAACTAGCTTTGTGAGAAATTTTG 660
Db 784 ATCCCGAGTCTCATTTCCGCGCTGCTGAGTAACTAGCTTTGTGAGAAATTTTG 843
Qy 661 CCGTGTGTGACAAATCAGTAGATTTTCATTGAGAAACAAAGCCTCAATGATGTT 720
Db 844 CCGTGTGTGACAAATCAGTAGATTTTCATTGAGAAACAAAGCCTCAATGATGTT 903
Qy 721 CTAGTGAACGTGTTAGCTGGGATTCCTCCGCTCCGCCAATCGCTATCGCTACATCATG 780
Db 904 CTAGTGAACGTGTTAGCTGGGATTCCTCCGCTCCGCCAATCGCTATCGCTACATCATG 963
Qy 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAAACCTACT 840
Db 964 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAAACCTACT 1023
Qy 841 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACTATGAGAAAGATTAGAAC 900
Db 1024 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACTATGAGAAAGATTAGAAC 1083
Qy 901 CAGACTGAGACATCAGGGCCAAAGAGCAACTCAAGCTGTGCACTGAGAAAGCCAAAT 960
Db 1084 CAGACTGAGACATCAGGGCCAAAGAGCAACTCAAGCTGTGCACTGAGAAAGCCAAAT 1143
Qy 961 GAACCTGTCCCTGCTCTCAGAGGGTGAACAGAAAGAGAGCGCCCTCAGTCCACCC 1020
Db 1144 GAACCTGTCCCTGCTCTCAGAGGGTGAACAGAAAGAGAGCGCCCTCAGTCCACCC 1203
Qy 1021 TGTGCGAAGCTCTCTAATCTCAGAGGAGAGCAAAAGGCCGTGATCCGCCAGCGTG 1080
Db 1204 TGTGCGAAGCTCTCTAATCTCAGAGGAGAGCAAAAGGCCGTGATCCGCCAGCGTG 1263
Qy 1081 CCCAGGCTGCGCAGGCTGAGCGGTGCTGTTAGAGAGACAGCCGCTGTGTAACAGCGCTC 1140
Db 1264 CCCAGGCTGCGCAGGCTGAGCGGTGCTGTTAGAGAGACAGCCGCTGTGTAACAGCGCTC 1323
Qy 1141 AGTGGGCTGACCTGTCCGAGACAGGCTGAGAAAGAGCAATTAAGCTTAAGGTTCCCTC 1200
Db 1324 AGTGGGCTGACCTGTCCGAGACAGGCTGAGAAAGAGCAATTAAGCTTAAGGTTCCCTC 1383
Qy 1201 TCTCTGGAATCAATCAAGTTCAATTAATTCAGCCAGCATGCGCATCTTACATGCTTC 1260
Db 1384 TCTCTGGAATCAATCAAGTTCAATTAATTCAGCCAGCATGCGCATCTTACATGCTTC 1443
Qy 1261 TCTCATCAGAAAGATGCTTTGAATTAACAACCTTCACTACTGTGATGGAGCAAC 1320
PR |||||

Db 1444 TCTCATCAGAAAGATGCTTTGAATTAACAACCTTCACTACTGTGATGGAGCAAC 1503
Qy 1321 AAGCTATGCAAGTCTCTCCCTGTTGAGAACTATGAGAGCAGACTCCGAAACCAAGTCT 1380
Db 1504 AAGCTATGCAAGTCTCTCCCTGTTGAGAACTATGAGAGCAGACTCCGAAACCAAGTCT 1563
Qy 1381 GATTAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGACAGCCGCAAGCTTACAGACGAG 1440
Db 1564 GATTAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGACAGCCGCAAGCTTACAGACGAG 1623
Qy 1441 AGCAAGCATTCATTCGCTCAGAAACAGCAGACAGTGGACCGCCAGAGGCTCCCTTTA 1500
Db 1624 AGCAAGCATTCATTCGCTCAGAAACAGCAGACAGTGGACCGCCAGAGGCTCCCTTTA 1683
Qy 1501 TCTTCACATGATCAGAAAGGAGCGTGAAGCAATTAACAACAGGCTTCTTTTCCG 1560
Db 1684 TCTTCACATGATCAGAAAGGAGCGTGAAGCAATTAACAACAGGCTTCTTTTCCG 1743
Qy 1561 CTTTCACACAGCAGACAGACCTCAAGAGTCTGCTGAGCTTGAAGGCTTGAC 1620
Db 1744 CTTTCACACAGCAGACAGACCTCAAGAGTCTGCTGAGCTTGAAGGCTTGAC 1803
Qy 1621 TCGGATATCTTGAGCCGCCAGACCTCTACCCCTTCCCTGACAGAGCTGGATTTG 1680
Db 1804 TCGGATATCTTGAGCCGCCAGACCTCTACCCCTTCCCTGACAGAGCTGGATTTG 1863
Qy 1681 ACAGAGTCTCACTTCTACTCTGCTCAGCCATCTACGAGGAGGAGTCCAGTTACTCT 1740
Db 1864 ACAGAGTCTCACTTCTACTCTGCTCAGCCATCTACGAGGAGGAGTCCAGTTACTCT 1923
Qy 1741 GCTTACAGCTGAGCAGCTGCGCACTTGAGGAGACCAAGCTATTCGTGAGGAGCGG 1800
Db 1924 GCTTACAGCTGAGCAGCTGCGCACTTGAGGAGACCAAGCTATTCGTGAGGAGCGG 1983
Qy 1801 CAGAAAGCCAAAGTGAACAGTCTGAGCGGAGAGCTGAGCAATGAAGAGAGCCCTTTGAA 1860
Db 1984 CAGAAAGCCAAAGTGAACAGTCTGAGCGGAGAGCTGAGCAATGAAGAGAGCCCTTTGAA 2043
Qy 1861 AAGCATTTAAACGCAAGCTGCGCAATGGAATTTGAGAGAGCATCATGTCAGAGAAC 1920
Db 2044 AAGCATTTAAACGCAAGCTGCGCAATGGAATTTGAGAGAGCATCATGTCAGAGAAC 2103
Qy 1921 AGTTCAGGAGAAAGCTGGGAAAGTGGGAGAGTGAAGTCTGTTTGGGAGAGATGAA 1980
Db 2104 AGTTCAGGAGAAAGCTGGGAAAGTGGGAGAGTGAAGTCTGTTTGGGAGAGATGAA 2163
Qy 1981 ATCATTTGAGTCTCC 1995
Db 2164 ATCATTTGAGTCTCC 2178

RESULT 9
ABV20833
ID ABV20833 standard; cDNA; 5145 BP.
XX AC ABV20833;
XX AC
XX ABV20833;
XX AC
XX 13-SEP-2002 (first entry)
XX DT
XX Human prostate expression marker cDNA 20824.
XX DE
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
```

PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 3419; 11750DP; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABN00010-ABN62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SO Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 94.9%; Score 1893; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGSGCCATGAGATGATGGAATGCAATGTTGTAAGAGGTTGGTGTCTGTGGA	60
DB	589	ATGSCCATGAGATGATGGAATGCAATGTTGTAAGAGGTTGGTGTCTGTGGA	648
QY	61	AGTGAACGGAAGAGTGTCTAATGATAGCCGCGCATTTGTGAAATACATATCC	120
DB	649	AGTGAACGGAAGAGTGTCTAATGATAGCCGCGCATTTGTGAAATACATATCC	708
QY	121	CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAAGCGAGTTGCAAC	180
DB	709	CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAAGCGAGTTGCAAC	768
QY	181	GACAAAGTGTATATACAGAGCTCATTCAGCATTCAGCGAAACATAAGTTGACAT	240
DB	769	GACAAAGTGTATATACAGAGCTCATTCAGCATTCAGCGAAACATAAGTTGACAT	828
QY	241	TGCAATCAGAAAGTGTATATACAGATCAAGCTCCCAAGATGTTGGCTCTCTCT	300
DB	829	TGCAATCAGAAAGTGTATATACAGATCAAGCTCCCAAGATGTTGGCTCTCTCT	888
QY	301	GACTGTTTCTCACTGACTTCTGGGTAACCTGAGAAAGCTTCAACTCTGTTCA	360
DB	889	GACTGTTTCTCACTGACTTCTGGGTAACCTGAGAAAGCTTCAACTCTGTTCA	948
QY	361	CTTGCAAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCCTCTGTGAAAGAA	420
DB	949	CTTGCAAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCCTCTGTGAAAGAA	1008
QY	421	ACTCTAGTCCCTACTGATCTTCTCAGCCTTGTGCTTATGTCGAACATTTGG	480
DB	1009	ACTCTAGTCCCTACTGATCTTCTCAGCCTTGTGCTTATGTCGAACATTTGG	1068
QY	481	CGAATTTCTTCCCAATCTTTATCTTGGCTGCAGGAGATGCTCTCAACGAAG	540
DB	1069	CGAATTTCTTCCCAATCTTTATCTTGGCTGCAGGAGATGCTCTCAACGAAG	1128

QY	541	CAGCAGAAATGGATGGTATATGTTAAATGACAGCTATATCTGTCCAAAGCTGACT	600
DB	1129	CAGCAGAAATGGATGGTATATGTTAAATGACAGAAATACCTGTCCAAAGCTGACT	1188
QY	601	ATCCCGAGTCTCATTTCTCGTGTGCTGTGATGACAGCTTTTGTGAAAAATTTTG	660
DB	1189	ATCCCGAGTCTCATTTCTCGTGTGCTGTGATGACAGCTTTTGTGAAAAATTTTG	1248
QY	661	CCGTGTTGAGCAAAATCACTATATTTGATGAAAGCAAAAGCCCTCCAAATGATG	720
DB	1249	CCGTGTTGAGCAAAATCACTATATTTGATGAAAGCAAAAGCCCTCCAAATGATG	1308
QY	721	CTAGTCACTGTTTATGCTGGATCTCCGCTCGGACCATGCTATTCCTTACATATG	780
DB	1309	CTAGTCACTGTTTATGCTGGATCTCCGCTCGGACCATGCTATTCCTTACATATG	1368
QY	781	AAGAGATGAGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAGAACT	840
DB	1369	AAGAGATGAGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAGAACT	1428
QY	841	ATATCTCCAAATTTCTTCTGCGGCAATCTCTGACTATGAAAGAAAGATTAAAG	900
DB	1429	ATATCTCCAAATTTCTTCTGCGGCAATCTCTGACTATGAAAGAAAGATTAAAG	1488
QY	901	CAGACTGAGCATCAGGCGCAAGCAAACTCAAGCTGCTGCACTGGAAGCAAAAT	960
DB	1489	CAGACTGAGCATCAGGCGCAAGCAAACTCAAGCTGCTGCACTGGAAGCAAAAT	1548
QY	961	GAACTGTCTCTGCTGTCTCAGAGGTGACAGAAAGCGAGAGCGCCCTCACTGCA	1020
DB	1549	GAACTGTCTCTGCTGTCTCAGAGGTGACAGAAAGCGAGAGCGCCCTCACTGCA	1608
QY	1021	TGTGCGGATCTGCTATCTCAGAGGACAGAGCAAAAGCCGCTGATCCCGACGCT	1080
DB	1609	TGTGCGGATCTGCTATCTCAGAGGACAGAGCAAAAGCCGCTGATCCCGACGCT	1668
QY	1081	CCGAGCGTCCAGCGGTGACGCGTGTGTAGAGACAGCCGCTGTGATCAGGCGCT	1140
DB	1669	CCGAGCGTCCAGCGGTGACGCGTGTGTAGAGACAGCCGCTGTGATCAGGCGCT	1728
QY	1141	AGTGGCTGACCTGTCTCCAGACAGAGCTGGAAGACAGCAATTAAGCTCAAGCT	1200
DB	1729	AGTGGCTGACCTGTCTCCAGACAGAGCTGGAAGACAGCAATTAAGCTCAAGCT	1788
QY	1201	TCTCTGATATCAATCACTGTTATATTCAGCCAGATGAGCATCTTATCATGCTTC	1260
DB	1789	TCTCTGATATCAATCACTGTTATATTCAGCCAGATGAGCATCTTATCATGCTTC	1848
QY	1261	TCCTCATGAAAGATGCTTGGAAATCTACAAACCTTCCACTCTGATGAGGACCA	1320
DB	1849	TCCTCATGAAAGATGCTTGGAAATCTACAAACCTTCCACTCTGATGAGGACCA	1908
QY	1321	AGCTATGCAAGTCTCTCCCTGTTTCAAGAACTATCGAGACAGCTCCGAAACAG	1380
DB	1909	AGCTATGCAAGTCTCTCCCTGTTTCAAGAACTATCGAGACAGCTCCGAAACAG	1968
QY	1381	GATAGGAGAAAGCCAGCATCTCCCAAGAAAGCTGCAACCGCCAGGCTTCAAG	1440
DB	1969	GATAGGAGAAAGCCAGCATCTCCCAAGAAAGCTGCAACCGCCAGGCTTCAAG	2028
QY	1441	AGCAAGCATGATGATTCGATTCAGAACACAGACAGTGGGACCGCCAGAGGTC	1500
DB	2029	AGCAAGCATGATGATTCGATTCAGAACACAGACAGTGGGACCGCCAGAGGTC	2088
QY	1501	TCCTCACTGATCGAAGTGGAGACGTGAGACAAATTAACACACAGCTTCTTTT	1560
DB	2089	TCCTCACTGATCGAAGTGGAGACGTGAGACAAATTAACACACAGCTTCTTTT	2148
QY	1561	CTTTCAACAGGACAGACAGCTCAGAGAGTCTGCTGCTGGGCTTTAAGGCTG	1620
DB	2149	CTTTCAACAGGACAGACAGCTCAGAGAGTCTGCTGCTGGGCTTTAAGGCTG	2208
QY	1621	TGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGATTTTGGC	1680

Db 2209 TCGATATCTTGCCCCCAGACCTCTACCCCTTCCCTGACACAGCTGTATTTTCC 2268
Qy 1661 ACAGAGTCTTCACTTCTACTCTGCTCAGCCATCTTAACGAGGCACTGCCAGTTACTCT 1740
Db 2269 ACAGAGTCTTCACTTCTACTCTGCTCAGCCATCTTAACGAGGCACTGCCAGTTACTCT 2228
Qy 1741 GCTTACAGCTGACGAGCTGCCCACTTGGCGAGACCAAGCTTATTTCTGTGCGGAGCGG 1800
Db 2229 GCTTACAGCTGACGAGCTGCCCACTTGGCGAGACCAAGCTTATTTCTGTGCGGAGCGG 2288
Qy 1801 CAGAGCCAAAGTGAACAGAGCTACTCTGCGGCGAGCTGAGCATGAAGAGAGCCCTTTGAA 1860
Db 2389 CAGAGCCAAAGTGAACAGAGCTACTCTGCGGCGAGCTGAGCATGAAGAGAGCCCTTTGAA 2448
Qy 1861 AAGCAGTTTAAACCCAGAGAGCTGCCAAATGTAATTTGAGAGAGCATCATGTCAAGAAC 1920
Db 2449 AAGCAGTTTAAACCCAGAGAGCTGCCAAATGTAATTTGAGAGAGCATCATGTCAAGAAC 2508
Qy 1921 AGCTCAGCGGAAAGCTGGGGAAAGTGGGAGTCAAGCTTATTTCTGCGGCGAGCATGGA 1980
Db 2509 AGCTCAGCGGAAAGCTGGGGAAAGTGGGAGTCAAGCTTATTTCTGCGGCGAGCATGGA 2568
Qy 1981 ATCAATTGAGGTCTCC 1995
Db 2569 ATCAATTGAGGTCTCC 2583

RESULT 10

ABV21080

ID ABV21080 standard; cDNA; 5145 BP.

XX AC ABV21080;

DT 13-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 21071.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

XX (MIL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of

XX PT prostate cells and correlating with presence of prostate cancer, useful

XX PT for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 3481; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX specification or its complement. (I) is useful for: (a) assessing whether

XX a patient is afflicted with prostate cancer; (b) monitoring the

XX progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 94.9%; Score 1893; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATTTGGAACCTCAATTTGTAATGAGAGGTGTGCTCTGCTGGA 60
Db 589 ATGGCCCATGAGATGATTTGGAACCTCAATTTGTAATGAGAGGTGTGCTCTGCTGGA 648
Qy 61 AGTGAACGGAAAAAGTGTCTGTAATTTGATAGCCGGCCATTTTGTGGAATTAATATCATCC 120
Db 649 AGTGAACGGAAAAAGTGTCTGTAATTTGATAGCCGGCCATTTTGTGGAATTAATATCATCC 708
Qy 121 CACATTTTGAAGCATTAAATATCAAGTCTCCAAAGTTTATGAAGCGAAGTTGCAACAG 180
Db 709 CACATTTTGAAGCATTAAATATCAAGTCTCCAAAGTTTATGAAGCGAAGTTGCAACAG 768
Qy 181 GACAAAGTGTATTAATTAAGAGCTCATCCAGCATTTGAGGAAACATTAAGTTGACATTGAT 240
Db 769 GACAAAGTGTATTAATTAAGAGCTCATCCAGCATTTGAGGAAACATTAAGTTGACATTGAT 828
Qy 241 TGCAGTCAAGAGGTGTATGTTTATGATCAATGAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
Db 829 TGCAGTCAAGAGGTGTATGTTTATGATCAATGAAGCTCCCAAGATGTTGCTCTCTCTTCA 888
Qy 301 GACTGTTTCTCACTGTACTTCTGGGTAACTGGAAGAGACTTCAACTCTGTTCAACCTG 360
Db 889 GACTGTTTCTCACTGTACTTCTGGGTAACTGGAAGAGACTTCAACTCTGTTCAACCTG 948
Qy 361 CTTCAGAGTGGGTTTGTGAGTCTCTCTGTTGTTTCCCTGCTCTGTGAAGAAATTC 420
Db 949 CTTCAGAGTGGGTTTGTGAGTCTCTCTGTTGTTTCCCTGCTCTGTGAAGAAATTC 1008
Qy 421 ACTGAGTCCCTACCTGATTTCTCAGCTTTGCTTACCTGTTGCAACATTTGGGCCAACC 480
Db 1009 ACTGAGTCCCTACCTGATTTCTCAGCTTTGCTTACCTGTTGCAACATTTGGGCCAACC 1068
Qy 481 CGAATTTTCCCAATCTTTATCTTGCTGCGCAGCAGAGATGCTCTCAACAGAGCTGATATG 540
Db 1069 CGAATTTTCCCAATCTTTATCTTGCTGCGCAGCAGAGATGCTCTCAACAGAGCTGATATG 1128
Qy 541 CAGCAGAAATGGGATTTGTTATGTTAAATGCCAGTAACTGTGTCGAAAGCTGACTTT 600
Db 1129 CAGCAGAAATGGGATTTGTTATGTTAAATGCCAGTAACTGTGTCGAAAGCTGACTTT 1188
Qy 601 ATCCCGAGTCTCAATTTCCGCGTGTGCTGTGAATACGCTTTTGTGAAGAAATTTTG 660
Db 1189 ATCCCGAGTCTCAATTTCCGCGTGTGCTGTGAATACGCTTTTGTGAAGAAATTTTG 1248
Qy 661 CCGTGTGTTGACCAATAGATGATTTTCAATTTGAGAAACCAAAAGCTTCAATGATGTGT 720
Db 1249 CCGTGTGTTGACCAATAGATGATTTTCAATTTGAGAAACCAAAAGCTTCAATGATGTGT 1308
Qy 721 CTATGTGACGTGTTAGCTGGGATTTCCCGCTCGGCCACCATGCTATGCTTACATCATG 780
Db 1309 CTATGTGACGTGTTAGCTGGGATTTCCCGCTCGGCCACCATGCTATGCTTACATCATG 1368
Qy 781 AAGAGATGACATGCTTTAGATGAGAGCTTACAGATTTTGAAGAAAAAAGCTACT 840
Db 1369 AAGAGATGACATGCTTTAGATGAGAGCTTACAGATTTTGAAGAAAAAAGCTACT 1428
Qy 841 ATATCTCAAACTTCAATTTTCTGGGCCCACTCTGAGCATATGAGAAAGATTAAGAAC 900
Db 1429 ATATCTCAAACTTCAATTTTCTGGGCCCACTCTGAGCATATGAGAAAGATTAAGAAC 1488

Db 709 CACATTTTGAAGCCATTAAATCAACTCTCCAACTTATGAAAGCAAGTTGCAACG 768
Qy 181 GCAAGGTGTAATACAGAGCTCAATCCAGCAATTCAGCAAAATAGGTTGACATTGAT 240
Db 769 GACAAAGTGTAAATACAGAGCTCAATCCAGCAATTCAGCAAAATAGGTTGACATTGAT 828
Qy 241 TSCAGTCAGAAAGTTGATGATTAAGATCAAGTCCCAAGATGTTGCTCTCTCTTCA 300
Db 829 TSCAGTCAGAAAGTTGATGATTAAGATCAAGTCCCAAGATGTTGCTCTCTCTTCA 888
Qy 301 GACTGTTTCTACTGTAATCTTGAGTAACTGAGAAAGAGCTTCACTGTTCACTG 360
Db 889 GACTGTTTCTACTGTAATCTTGAGTAACTGAGAAAGAGCTTCACTGTTCACTG 948
Qy 361 CTTCGAGGTGGTTGCTGAGTCTCTGTTGTTTCTCTGAGCTCTGTAAGAAATCC 420
Db 949 CTTCGAGGTGGTTGCTGAGTCTCTGTTGTTTCTCTGAGCTCTGTAAGAAATCC 1008
Qy 421 ACTTAGTCCCTACCTGATTTCTCAGGCTTGTACCTGTTGCAAACTTGGGCAAC 480
Db 1009 ACTTAGTCCCTACCTGATTTCTCAGGCTTGTACCTGTTGCAAACTTGGGCAAC 1068
Qy 481 CGAATTTCTCCCAATCTTATCTTGAGTAACTGAGTAACTGAGTAACTGAGTAA 540
Db 1069 CGAATTTCTCCCAATCTTATCTTGAGTAACTGAGTAACTGAGTAACTGAGTAA 1128
Qy 541 CAGAGATGGGATGTTGTTATGTTAAATGCAAGTAACTGAGTAACTGAGTAACTG 600
Db 1129 CAGAGATGGGATGTTGTTATGTTAAATGCAAGTAACTGAGTAACTGAGTAACTG 1188
Qy 601 ATCCCGAGTCTCATTTCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTG 660
Db 1189 ATCCCGAGTCTCATTTCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTG 1248
Qy 661 CCGTGTGTCAGCAATAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTG 720
Db 1249 CCGTGTGTCAGCAATAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTG 1308
Qy 721 CTAGTGCATGTTTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAA 780
Db 1309 CTAGTGCATGTTTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAA 1368
Qy 781 AAGAGATGACATGCTTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTG 840
Db 1369 AAGAGATGACATGCTTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTG 1428
Qy 841 ATATCTCCAACTTCAATTTCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTG 900
Db 1429 ATATCTCCAACTTCAATTTCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTG 1488
Qy 901 CAACTGAGCATCAGGCGCAAAAGCAAACTCAGCTGCTGCACTGAGAGAGCAAT 960
Db 1489 CAACTGAGCATCAGGCGCAAAAGCAAACTCAGCTGCTGCACTGAGAGAGCAAT 1548
Qy 961 GAACTGTCCCTGCTCTCAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 1549 GAACTGTCCCTGCTCTCAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1608
Qy 1021 TGTGCGCACTGCTAATCTCAGAGGCAAGAGCAAAAGCGCTGATCCGCGCAGCTG 1080
Db 1609 TGTGCGCACTGCTAATCTCAGAGGCAAGAGCAAAAGCGCTGATCCGCGCAGCTG 1668
Qy 1081 CCAGGCTGCGCAAGCTGAGAGCTGCTGTTAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1669 CCAGGCTGCGCAAGCTGAGAGCTGCTGTTAGAGAGAGAGAGAGAGAGAGAGAGAG 1728
Qy 1141 AGTGGCTGCACTGCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1729 AGTGGCTGCACTGCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1788
Qy 1201 TCTCTGATATCAATCAGTTTCAATTTCAAGCAAGTGAAGTCTTCAATGAGCTTC 1260

Db 1789 TCTCTGATATCAATCAGTTTCAATTTCAATTCAGCCAGATGAGAGATCTTACATGGCTTC 1848
Qy 1261 TCCCTATCAGAAAGCTTTGGAATCTACAAACCTTCCACTACTTGTGATGGAGCAAC 1320
Db 1849 TCCCTATCAGAAAGCTTTGGAATCTACAAACCTTCCACTACTTGTGATGGAGCAAC 1908
Qy 1321 AAGCTATGCCAGTTCTCCCTGTTTCAAGAACTATCGAGACAGACTCCGAAACAGTCT 1380
Db 1909 AAGCTATGCCAGTTCTCCCTGTTTCAAGAACTATCGAGACAGACTCCGAAACAGTCT 1968
Qy 1381 GATTAAG 1440
Db 1969 GATTAAG 2028
Qy 1441 AGCAAGCATGCTGCTGAG 1500
Db 2029 AGCAAGCATGCTGCTGAG 2088
Qy 1501 TCTCCACTGATGCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 2089 TCTCCACTGATGCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2148
Qy 1561 CTTTCCAGCAGCAG 1620
Db 2149 CTTTCCAGCAGCAG 2208
Qy 1621 TCGATATCTTGGCCCCCAG 1680
Db 2209 TCGATATCTTGGCCCCCAG 2268
Qy 1681 ACAGAGTCTCACTTCTACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 2269 ACAGAGTCTCACTTCTACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2328
Qy 1741 GCTTACAGCTCAGCAG 1800
Db 2329 GCTTACAGCTCAGCAG 2388
Qy 1801 CAG 1860
Db 2389 CAG 2448
Qy 1861 AAGAGATTTAAACAG 1920
Db 2449 AAGAGATTTAAACAG 2508
Qy 1921 AGGTACAGGAG 1980
Db 2509 AGGTACAGGAG 2568
Qy 1981 ATCATTTGAGGTTCC 1995
Db 2569 ATCATTTGAGGTTCC 2583

RESULT 12
ABV20978
ID ABV20978 standard; cDNA; 5145 BP.
XX ABV20978;
XX
XX 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 20969.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX
XX W0200160860-A2.
XX
XX
XX
XX PD 23- AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.
 XX 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Schlegel R, Endege WO, Monahan JR;
 XX WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 3451; 11750p; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (1) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (1) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
 CC
 XX
 SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
 Query Match 94.9%; Score 1893; DB 5; Length 5145;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 481 CGAATCTTCCCAATCTTATCTGGCTGCACAGAGATGTCTCTCAACAAGAGCTGATA 540
 |||||
 DB 1069 CGAATCTTCCCAATCTTATCTGGCTGCACAGAGATGTCTCTCAACAAGAGCTGATG 1128
 QY 541 CAGCAGAAATGGGATTTGTTATGTTAAATGCCAGTATACCTGTCCAAAGCTGACTTT 600
 |||||
 DB 1129 CAGCAGAAATGGGATTTGTTATGTTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT 1188
 QY 601 ATCCCGAGTCTCATTTCTCGGTGTGCTGTGTAATGACAGCTTTTGTGAGAAATTTTG 660
 |||||
 DB 1189 ATCCCGAGTCTCATTTCTCGGTGTGCTGTGTAATGACAGCTTTTGTGAGAAATTTTG 1248
 QY 661 CCGTGTGAGCAAAATCAGATGATTTCTATTGAGAAAGCAAAAGCTTCATGATGATGTT 720
 |||||
 DB 1249 CCGTGTGAGCAAAATCAGATGATTTCTATTGAGAAAGCAAAAGCTTCATGATGATGTT 1308
 QY 721 CTAGTCACTGTTTACCTGGGATCTCCGCTCCGCCACCATGCTATGCTTACATCATG 780
 |||||
 DB 1309 CTAGTCACTGTTTACCTGGGATCTCCGCTCCGCCACCATGCTATGCTTACATCATG 1368
 QY 781 AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGAAAGCTACT 840
 |||||
 DB 1369 AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGAAAGCTACT 1428
 QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGAAATTAAGAC 900
 |||||
 DB 1429 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGAGCTATGAGAAAGAAATTAAGAC 1488
 QY 901 CAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGACCTGAGAGAGCCAAAT 960
 |||||
 DB 1489 CAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGACCTGAGAGAGCCAAAT 1548
 QY 961 GAACTGTCTCTGTCTCTCAAGAGGTGACAGAAAGAGAGAGCCCTCACTCAACCC 1020
 |||||
 DB 1549 GAACTGTCTCTGTCTCTCAAGAGGTGACAGAAAGAGAGAGCCCTCACTCAACCC 1608
 QY 1021 TGTGCCGACTGTGCTACCTCAGAGGACAGAGCAAAAGCCGTGATCCGCAAGCTG 1080
 |||||
 DB 1609 TGTGCCGACTGTGCTACCTCAGAGGACAGAGCAAAAGCCGTGATCCGCAAGCTG 1668
 QY 1081 CCCAGGCTCCAGAGCTGTGAGGCTGTCTGTTAGAGACAGCCGCTGTGATCAGGCGCT 1140
 |||||
 DB 1669 CCCAGGCTCCAGAGCTGTGAGGCTGTCTGTTAGAGACAGCCGCTGTGATCAGGCGCT 1728
 QY 1141 AGTGGCTGCACTGTCTGCGCAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTCTTC 1200
 |||||
 DB 1729 AGTGGCTGCACTGTCTGCGCAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTCTTC 1788
 QY 1201 TCTCTGATATCAAAATCACTTTTATATTCAGACAGATGAGAGATCTTACATGCTTC 1260
 |||||
 DB 1789 TCTCTGATATCAAAATCACTTTTATATTCAGACAGATGAGAGATCTTACATGCTTC 1848
 QY 1261 TCTCTATCAGAAAGATCTTTGGAATCTCAAACTTCCATCACTCTGATGGAGCAAC 1320
 |||||
 DB 1849 TCTCTATCAGAAAGATCTTTGGAATCTCAAACTTCCATCACTCTGATGGAGCAAC 1908
 QY 1321 AAGCTATGCAAGTTCCTCCCTGTTCAAGAACTATGAGACAGCTCCGAAACCAAGTCT 1380
 |||||
 DB 1909 AAGCTATGCAAGTTCCTCCCTGTTCAAGAACTATGAGACAGCTCCGAAACCAAGTCT 1968
 QY 1381 GATTAAGAGAAAGCCAGCATCTCCCAAGAGCTGACAGCCGAGGCTTCAAGACCCAG 1440
 |||||
 DB 1969 GATTAAGAGAAAGCCAGCATCTCCCAAGAGCTGACAGCCGAGGCTTCAAGACCCAG 2028
 QY 1441 AGCAAGCAATTCATTCGCTCAGAAACCAAGCAGAGTGGACCGCCAGAGGCTCCCTTTA 1500
 |||||
 DB 2029 AGCAAGCAATTCATTCGCTCAGAAACCAAGCAGAGTGGACCGCCAGAGGCTCCCTTTA 2088
 QY 1501 TCTCCACTGATCGAAGTGGAGCGGTGAGAGCAATTAACCAACAGCTCTCTTTTGGC 1560
 |||||
 DB 2089 TCTCCACTGATCGAAGTGGAGCGGTGAGAGCAATTAACCAACAGCTCTCTTTTGGC 2148
 QY 1561 CTTTCCACAGCAGAGACAGCACTCAAGAGTCTGCTGGGCTTAAAGGCTTGCAAC 1620

DB 2149 CTTTCACACGACGACGACCTCAGAGAGTCTGCGCTGGCCTTAAGGCTGGAC 2208
QY 1621 TCGGATATCTTGGCCCCCGACGACCTTCCCTGACGACGCTGTAATTTTGC 1680
DB 2209 TCGGATATCTTGGCCCCCGACGACCTTCCCTGACGACGCTGTAATTTTGC 2268
QY 1681 ACAGAGTCTCAGACTTCTACTTGTGCTCAGGCACTTAAGGAGGCTGCCAGTTACT 1740
DB 2269 ACAGAGTCTCAGACTTCTACTTGTGCTCAGGCACTTAAGGAGGCTGCCAGTTACT 2328
QY 1741 GCCTACAGCTGACGACGACCTGCGCCTGCGAGACCAAGTCTATCTGTGCGACGCG 1800
DB 2329 GCCTACAGCTGACGACGACCTGCGCCTGCGAGACCAAGTCTATCTGTGCGACGCG 2388
QY 1801 CAGAGCGCAGTACGACGACGCTGCTGCGCGGAGCTGGCATGAGAGCCCTTTGAA 1860
DB 2389 CAGAGCGCAGTACGACGACGCTGCTGCGCGGAGCTGGCATGAGAGCCCTTTGAA 2448
QY 1861 AAGCAGTTTAAACGACGACGCTGCCAATGGAATTTGGAGAGCATCTGTCAAGAAC 1920
DB 2449 AAGCAGTTTAAACGACGACGCTGCCAATGGAATTTGGAGAGCATCTGTCAAGAAC 2508
QY 1921 AGGTCACGGAGAGCTGGGGAAGTGGGAGTCACTTGTGCGGACAGTGA 1980
DB 2509 AGGTCACGGAGAGCTGGGGAAGTGGGAGTCACTTGTGCGGACAGTGA 2568
QY 1981 ATCATTGAGGTCTCC 1995
DB 2569 ATCATTGAGGTCTCC 2583

RESULT 13

ABV21092 ID ABV21092 standard; cDNA; 5145 BP.

XX AC ABV21092;

XX DT 13-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 21083.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN W0200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAY-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUL-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX PS Claim 1; Page 3485; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the CC specification or its complement. (I) is useful for: (a) assessing whether CC a patient is afflicted with prostate cancer; (b) monitoring the CC progression of prostate cancer in a patient; (c) assessing the efficacy CC of a test compound to inhibit prostate cancer in a patient; (d) assessing CC the efficacy of a therapy for inhibiting prostate cancer in a patient; CC (e) selecting a composition for inhibiting prostate cancer in a patient; CC (f) assessing the prostate cell carcinogenic potential of a compound; (g) CC determining whether prostate cancer has metastasized in a patient; (h) CC assessing the aggressiveness or incidence of prostate cancer in a patient CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 94.9%; Score 1893; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAACCTCAATTGTTACTGAGAGGTGGTGGCTCTGAGAA 60
DB 589 ATGGCCCATGAGATGATGGAACCTCAATTGTTACTGAGAGGTGGTGGCTCTGAGAA 648
QY 61 AGTGGACGGAAGAAAGTCTGCTAATGATGATGCGGCGCATTTTGGAAATACATATCCTC 120
DB 649 AGTGGACGGAAGAAAGTCTGCTAATGATGATGCGGCGCATTTTGGAAATACATATCCTC 708
QY 121 CACATTTTGAAGCCATTAATATCACTGCTCCAGCTTATGAGAGCGAAGGTTGCAACAG 180
DB 709 CACATTTTGAAGCCATTAATATCACTGCTCCAGCTTATGAGAGCGAAGGTTGCAACAG 768
QY 181 GACAAAGTGTAAATATACAGAGCTCATCCAGATTCAGGAAACATTAAGCTTGAATTGAT 240
DB 769 GACAAAGTGTAAATATACAGAGCTCATCCAGATTCAGGAAACATTAAGCTTGAATTGAT 828
QY 241 TGCAGTCAAGAGCTGTGATTTAGATTAAGAGCTCCCAAGATGTTGGCTCTGCTTCA 300
DB 829 TGCAGTCAAGAGCTGTGATTTAGATTAAGAGCTCCCAAGATGTTGGCTCTGCTTCA 888
QY 301 GACTGTTTCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 889 GACTGTTTCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
QY 361 CTGGCAGGTGGGTTTGTGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 949 CTGGCAGGTGGGTTTGTGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008
QY 421 ACTCTAGTCCCTCACTGATTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAG 480
DB 1009 ACTCTAGTCCCTCACTGATTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAG 1068
QY 481 CGAATTTCTCCCAATCTTATCTTGGCTGCCAGAGATGCTTCAACAAGAGCTGATA 540
DB 1069 CGAATTTCTCCCAATCTTATCTTGGCTGCCAGAGATGCTTCAACAAGAGCTGATA 1128
QY 541 CAGCAGAAATGGAGTTGTTATGTTAAATGCGACCTATCTGCTCCAAAGCTTGAATTT 600
DB 1129 CAGCAGAAATGGAGTTGTTATGTTAAATGCGACCTATCTGCTCCAAAGCTTGAATTT 1188
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 1189 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1248
QY 661 CCGTGGTTGGACAAATCAGTATGATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTGT 720
DB 1249 CCGTGGTTGGACAAATCAGTATGATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTGT 1308
QY 721 CTATGACACTGTTTATGCTGGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 780
DB 1309 CTATGACACTGTTTATGCTGGGATCTCCGCTCCGACCATGCTATGCTTACATCATG 1368
QY 781 AAGAGATGAGCATGCTTATGATGAAGCTTCAAGTTTGTGAAGAAAGAAAGAAAGCTTACT 840
DB 1369 AAGAGATGAGCATGCTTATGATGAAGCTTCAAGTTTGTGAAGAAAGAAAGAAAGCTTACT 1428

Db 649 AGTGAACGGAAAAAGTGGCTGAATTGATGAGCCGGCCATTTGGAAATACAAATCAATCC 708
Qy 121 CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAAGGAAAGTTGGAACG 180
Db 709 CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAAGGAAAGTTGGAACG 768
Qy 181 GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCAAAATATAGGTTGACATTAAT 240
Db 769 GACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCAAAATATAGGTTGACATTAAT 828
Qy 241 TGCAGTCAGAAAGTTGATTAGATCAATCAAAAGTCCCAAGATGTTGCTCTCTCTCA 300
Db 829 TGCAGTCAGAAAGTTGATTAGATCAATCAAAAGTCCCAAGATGTTGCTCTCTCTCA 888
Qy 301 GACTGTTTCTCATCTGACTTCTGGGTTAACTGAGAAAGAGCTTCAACTCTGTTCACTG 360
Db 889 GACTGTTTCTCATCTGACTTCTGGGTTAACTGAGAAAGAGCTTCAACTCTGTTCACTG 948
Qy 361 CTTCAGAGTGGGTTTGTGCTGAGTTCTCTGTTGTTTCTCTGGCTCTGTGAAGAAATCC 420
Db 949 CTTCAGAGTGGGTTTGTGCTGAGTTCTCTGTTGTTTCTCTGGCTCTGTGAAGAAATCC 1008
Qy 421 ACTCTAGTCCCTTACCTGCAATTTCTCAGCCTTGTCTTACCTGTTGCCAAATGGGCGCAAC 480
Db 1009 ACTCTAGTCCCTTACCTGCAATTTCTCAGCCTTGTCTTACCTGTTGCCAAATGGGCGCAAC 1068
Qy 481 CGAATTCCTCCCAATCTTTATCTTGGCTGCACAGAGATGTCCTCAACAAGAGCTGATA 540
Db 1069 CGAATTCCTCCCAATCTTTATCTTGGCTGCACAGAGATGTCCTCAACAAGAGCTGATA 1128
Qy 541 CAGCAGATGGGATTTGTTATGTTTAATGCCAGCTATACCTGTCCAAAGCTGACTT 600
Db 1129 CAGCAGATGGGATTTGTTATGTTTAATGCCAGCTATACCTGTCCAAAGCTGACTT 1188
Qy 601 ATCCCGAGTCTCATTTCCGCGTGGCTGTGATGACAGCTTTGTGAGAAATTTTG 660
Db 1189 ATCCCGAGTCTCATTTCCGCGTGGCTGTGATGACAGCTTTGTGAGAAATTTTG 1248
Qy 661 CCGTGTGTGACCAATCAGTAGATTTTATGAGAAAGCAAAAGCTTCCAAATGATGTT 720
Db 1249 CCGTGTGTGACCAATCAGTAGATTTTATGAGAAAGCAAAAGCTTCCAAATGATGTT 1308
Qy 721 CTAATGCACTGTTAGCTGGGATCTCCGCTCCGCGCAACATCGCTATCGCTACATG 780
Db 1309 CTAATGCACTGTTAGCTGGGATCTCCGCTCCGCGCAACATCGCTATCGCTACATG 1368
Qy 781 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAAGACTACT 840
Db 1369 AAGAGATGACATGCTTTAGATGAAAGCTTACAGATTTGTGAAAGAAAAAGACTACT 1428
Qy 841 ATATCTCCAAATCTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTAAAG 900
Db 1429 ATATCTCCAAATCTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTAAAG 1488
Qy 901 CAGACTGAGGATCAGGGGCCAAAGAGCAAACTCAGAGCTGACCTGAGAAAGCCAAAT 960
Db 1489 CAGACTGAGGATCAGGGGCCAAAGAGCAAACTCAGAGCTGACCTGAGAAAGCCAAAT 1548
Qy 961 GAACTGTCCCTGTCTGTCTCAGAGGGTGAAGAGAAAGAGAGGCCCTTCACTGTCACC 1020
Db 1549 GAACTGTCCCTGTCTGTCTCAGAGGGTGAAGAGAAAGAGAGGCCCTTCACTGTCACC 1608
Qy 1021 TGTGCGGACTGTCTGCTACCTCAGAGGACAGAGCAAAAGCCCGTGCATCCCGCAGGGTG 1080
Db 1609 TGTGCGGACTGTCTGCTACCTCAGAGGACAGAGCAAAAGCCCGTGCATCCCGCAGGGTG 1668
Qy 1081 CCGAGCTGCCCCAGGCTGAGCCGCTGCTTTAGAGAGCAGCCGCTGGTACAGGCGCTC 1140
Db 1669 CCGAGCTGCCCCAGGCTGAGCCGCTGCTTTAGAGAGCAGCCGCTGGTACAGGCGCTC 1728
Qy 1141 AGTGGGCTGACCTGTCTGCGAGACAGGCTGGAAGACAGCAATTAAGCTTAACGTTCTTC 1200

Db 1729 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGGCTTCCTTC 1788
Qy 1201 TCTCTGATATCAAAATCATGTTTATATTTACATGCCAGATGGCAGATTCCTTACATGGCTTC 1260
Db 1789 TCTCTGATATCAAAATCATGTTTATATTTACATGCCAGATGGCAGATTCCTTACATGGCTTC 1848
Qy 1261 TCTCTCATCAAAAGATGTTTGGAAATGTAACAACTTCCACTCTCTGTGATGGGACCAAC 1320
Db 1849 TCTCTCATCAAAAGATGTTTGGAAATGTAACAACTTCCACTCTCTGTGATGGGACCAAC 1908
Qy 1321 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATGGAAGACATCTCCGAAACCAATCTCT 1380
Db 1909 AAGCTATGCAAGTTCTCCCTGTTCAAGAACTATGGAAGACATCTCCGAAACCAATCTCT 1968
Qy 1381 GATTAAGAGGAGGACACATTCCTCAAGAAAGCTGACAGCCGAGGCTTACAGACGACG 1440
Db 1969 GATTAAGAGGAGGACACATTCCTCAAGAAAGCTGACAGCCGAGGCTTACAGACGACG 2028
Qy 1441 AGCAAGGATTTGCAATGCGTCTCAGAACAGAGAGTGGCAACCGCCAGAGGTCCTTTTA 1500
Db 2029 AGCAAGGATTTGCAATGCGTCTCAGAACAGAGAGTGGCAACCGCCAGAGGTCCTTTTA 2088
Qy 1501 TCTCCACTGCATGCAAGTGGAGGCTGGAAGACATTAACACACAGCTTCCTTTGCGC 1560
Db 2089 TCTCCACTGCATGCAAGTGGAGGCTGGAAGACATTAACACACAGCTTCCTTTGCGC 2148
Qy 1561 CTTTCCACACAGGACAGACCTTCAGAGAGTCTGAGGCTTGAAGGCTGAGC 1620
Db 2149 CTTTCCACACAGGACAGACCTTCAGAGAGTCTGAGGCTTGAAGGCTGAGC 2208
Qy 1621 TCGATATCTTGGGCCCCCAGACCTTAACCCCTTCCCTGACACAGAGCTGGATTTTGCC 1680
Db 2209 TCGATATCTTGGGCCCCCAGACCTTAACCCCTTCCCTGACACAGAGCTGGATTTTGCC 2268
Qy 1681 ACAGATCTCTACACTTTTACTCTGCTCAGCCATCTACAGAGGACAGTACCTCT 1740
Db 2269 ACAGATCTCTACACTTTTACTCTGCTCAGCCATCTACAGAGGACAGTACCTCT 2328
Qy 1741 GCTTACAGCTGACGACCTCCCACTTGCAGAGACCAAGTCTATTTCTGTGCGAGCGG 1800
Db 2329 GCTTACAGCTGACGACCTCCCACTTGCAGAGACCAAGTCTATTTCTGTGCGAGCGG 2388
Qy 1801 CAGAAAGCCAGTGAAGCTGACTCTGCGGCTGAGAGCTGAGTGAAGAGAGCCCTTGAA 1860
Db 2389 CAGAAAGCCAGTGAAGCTGACTCTGCGGCTGAGAGCTGAGTGAAGAGAGCCCTTGAA 2448
Qy 1861 AAGCAGTTTAAACGAGAAAGCTGCCAAATGGAATTTGAGAGAGATCATGTGAGAGAC 1920
Db 2449 AAGCAGTTTAAACGAGAAAGCTGCCAAATGGAATTTGAGAGAGATCATGTGAGAGAC 2508
Qy 1921 AGGTACCGGAAAGAGCTGGGAAAGTGGGAGTCACTTCTTGGGACAGATGGA 1980
Db 2509 AGGTACCGGAAAGAGCTGGGAAAGTGGGAGTCACTTCTTGGGACAGATGGA 2568
Qy 1981 ATCATTTAGAGTCTCC 1995
Db 2569 ATCATTTAGAGTCTCC 2583

RESULT 15
ABV21316
ID ABV21316 standard, cDNA, 5145 BP.
XX
XX ABV21316;
XX
DT 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker CDNA 21307.
DE
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.


```
Db      2089  TCTCACTGCATCGAAGTGGAGGTGGAGACATTACACACCAGCTTCCTTTGCGC 2148
Qy      1561  CTTTCCACCCGACGACGACCTTCACGAAGTCTGCTGCGCTTAAAGGCTGGCAC 1620
Db      2149  CTTTCCACCGACGACGACCTTCACGAAGTCTGCTGCGCTTAAAGGCTGGCAC 2208
Qy      1621  TCGGATATCTTGGCCCCCGACGACCTTACCCCTTCCCTGACGACGCTGTAATTTGGC 1680
Db      2209  TCGGATATCTTGGCCCCCGACGACCTTACCCCTTCCCTGACGACGCTGTAATTTGGC 2268
Qy      1681  ACAGAGTCTCACAATTCTACTGCTCAGCCATCTACGAGGCAAGTGCAGTTACTCT 1740
Db      2269  ACAGAGTCTCACAATTCTACTGCTCAGCCATCTACGAGGCAAGTGCAGTTACTCT 2328
Qy      1741  GCCTACAGCTGCAGCCGACGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCCAGCGG 1800
Db      2329  GCCTACAGCTGCAGCCGACGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCCAGCGG 2388
Qy      1801  CAGAAGCCAAGTACAGAGCTGACTGCGCGCGAGCTGGCATGAGAGAGCCCTTTGAA 1860
Db      2389  CAGAAGCCAAGTACAGAGCTGACTGCGCGCGAGCTGGCATGAGAGAGCCCTTTGAA 2448
Qy      1861  AAGCAAGTTTAAACGACGAAGCTGCCAAATGGAATTTGAGAGACATCATGTCAAGAA 1920
Db      2449  AAGCAAGTTTAAACGACGAAGCTGCCAAATGGAATTTGAGAGACATCATGTCAAGAA 2508
Qy      1921  AGGTACGGGAGAGAGCTGGGGAAGTGGGAGTCACTTAAGCTTTTCGGGCAGCATGAA 1980
Db      2509  AGGTACGGGAGAGAGCTGGGGAAGTGGGAGTCACTTAAGCTTTTCGGGCAGCATGAA 2568
Qy      1981  ATCATTGAGGTCTCC 1995
Db      2569  ATCATTGAGGTCTCC 2583
```

Search completed: June 22, 2004, 03:44:31
Job time : 1146 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 22, 2004, 03:25:27 ; Search time 3255 Seconds

(Without alignments)
2807.764 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccatgagatgattg9.....tggcaatcatgaggtctcc 1995Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3017426 seqs, 2290544650 residues

Word size : 0

Total number of hits satisfying chosen parameters: 820462

Minimum DB seq length: 1419

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
14: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*
19: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1893	94.9	1998	9 US-09-816-494-3	Sequence 3, Appl
2	1893	94.9	1998	16 US-10-377-072-27	Sequence 27, Appl
3	1893	94.9	3059	17 US-10-257-026-1	Sequence 1, Appl
4	1893	94.9	3496	9 US-09-964-277-1	Sequence 1, Appl
5	1893	94.9	3544	9 US-09-816-494-1	Sequence 1, Appl
6	1893	94.9	3544	16 US-10-377-072-25	Sequence 25, Appl
7	1893	94.9	3625	13 US-10-425-114-26334	Sequence 26234, A
8	1893	94.9	3766	13 US-10-343-357-17	Sequence 17, Appl
9	1893	94.9	4790	17 US-10-648-593-115	Sequence 115, Appl
10	1849	92.7	2966	13 US-10-296-115-520	Sequence 520, Appl
11	1842	92.3	2732	13 US-10-168-506-2	Sequence 2, Appl
12	1791	89.8	2102	16 US-10-094-749-673	Sequence 673, Appl
13	1418	71.1	3332	9 US-09-964-277-20	Sequence 20, Appl
14	1233	61.8	2071	13 US-10-072-012-257	Sequence 257, Appl

15	1182	59.2	2200	13 US-10-072-012-255	Sequence 255, Appl
16	788	39.5	1916	16 US-10-108-260A-2429	Sequence 2429, Appl
17	19	1.0	1449	12 US-10-052-482-192	Sequence 192, Appl
18	19	1.0	1635	17 US-10-648-593-27	Sequence 27, Appl
19	19	1.0	1830	15 US-10-346-356-1	Sequence 1, Appl
20	19	1.0	1909	12 US-10-052-482-191	Sequence 191, Appl
21	19	1.0	1909	13 US-10-058-270A-129	Sequence 129, Appl
22	19	1.0	2039	13 US-10-072-012-265	Sequence 265, Appl
23	19	1.0	2081	16 US-10-242-355-1076	Sequence 1076, Appl
24	19	1.0	2170	13 US-10-158-844-312	Sequence 312, Appl
25	19	1.0	2227	15 US-10-007-926A-330	Sequence 330, Appl
26	19	1.0	2227	16 US-10-435-696-54	Sequence 54, Appl
27	19	1.0	2387	16 US-10-443-694-100	Sequence 100, Appl
28	19	1.0	2476	13 US-10-220-120-28	Sequence 28, Appl
29	19	1.0	2631	17 US-10-378-448-3	Sequence 3, Appl
30	19	1.0	2653	9 US-09-969-708-536	Sequence 536, Appl
31	19	1.0	2653	10 US-09-960-706-716	Sequence 716, Appl
32	19	1.0	2653	10 US-09-873-319-451	Sequence 451, Appl
33	19	1.0	2653	15 US-10-094-699-2	Sequence 2, Appl
34	19	1.0	2653	15 US-10-205-823-131	Sequence 131, Appl
35	19	1.0	2653	15 US-10-210-120-5	Sequence 5, Appl
36	19	1.0	2653	16 US-10-341-434-70	Sequence 70, Appl
37	19	1.0	2653	16 US-10-117-937-7	Sequence 7, Appl
38	19	1.0	2653	16 US-10-443-694-1	Sequence 1, Appl
39	19	1.0	2655	15 US-10-252-157-9	Sequence 9, Appl
40	19	1.0	2731	16 US-10-264-048-696	Sequence 696, Appl
41	19	1.0	3017	16 US-10-443-694-38	Sequence 39, Appl
42	19	1.0	3017	16 US-10-443-694-128	Sequence 128, Appl
43	19	1.0	3077	16 US-10-443-694-90	Sequence 90, Appl
44	19	1.0	3286	10 US-09-971-392-9	Sequence 9, Appl
45	19	1.0	15500	9 US-09-764-869-2109	Sequence 2109, Appl

ALIGNMENTS

RESULT 1
US-09-816-494-3
; Sequence 3, Application US/09816494
; Patent No. US2002034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-816-494-3

Query Match 94.9%; Score 1893; DB 9; Length 1998;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATTGGAACTCAATTGTTACTGAGAGGTTGGCTGCTGCGGAA 60
1 ATGGCCCATGAGATGATTGGAACTCAAAATTGTTACTGAGAGGTTGGCTGCTGCGGAA 60
DB 1 ATGGCCCATGAGATGATTGGAACTCAAAATTGTTACTGAGAGGTTGGCTGCTGCGGAA 60
QY 61 AGTGAAGCGAAGAAAGTGTCTGATTAATGATGAGCGGCAATTTGTGGAATPACATATC 120
61 AGTGAAGCGAAGAAAGTGTCTGATTAATGATGAGCGGCAATTTGTGGAATPACATATC 120
DB 61 AGTGAAGCGAAGAAAGTGTCTGATTAATGATGAGCGGCAATTTGTGGAATPACATATC 120
QY 121 CACATTTTGAAGCACTTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACAG 180
121 CACATTTTGAAGCACTTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACAG 180
DB 121 CACATTTTGAAGCACTTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACAG 180

QY 181 GACAAAGTGTAAATTCAGAGCTCAATCCAGATTCAGGCAACATTAAGTTGACTTAT 240
DB 181 GACAAAGTGTAAATTCAGAGCTCAATCCAGATTCAGGCAACATTAAGTTGACTTAT 240
QY 241 TGCAGTCAGAGGTTGATTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 300
DB 241 TGCAGTCAGAGGTTGATTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 300
QY 301 GACTGTTTCTCAGTACTTCTGGGTAACTGAGAAAGAGCTTCACTCTGTTCACTG 360
DB 301 GACTGTTTCTCAGTACTTCTGGGTAACTGAGAAAGAGCTTCACTCTGTTCACTG 360
QY 361 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
DB 361 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
QY 421 ACTCTAGTCCCTAAGCTGATTTCTGAGCTTGTACCTGTGTCCTGTTGCAATTTGGCCAA 480
DB 421 ACTCTAGTCCCTAAGCTGATTTCTGAGCTTGTACCTGTGTCCTGTTGCAATTTGGCCAA 480
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGGCCAGAGAGATGTCCTCAAGAGCTGATA 540
DB 481 CGAATTTCTCCCAATCTTTATCTTGGCTGGCCAGAGAGATGTCCTCAAGAGCTGATA 540
QY 541 CAGCAGAAATGAGATTTGTTATGTTAAATGCGAGCTATACTGTCCAAAGCTGACTT 600
DB 541 CAGCAGAAATGAGATTTGTTATGTTAAATGCGAGCAATCTGTCCAAAGCTGACTT 600
QY 601 ATCCCGAGTCTCAATTTCCGCGGTGCGCTGAAATGACAGCTTTGTGAAATTTTG 660
DB 601 ATCCCGAGTCTCAATTTCCGCGGTGCGCTGAAATGACAGCTTTGTGAAATTTTG 660
QY 661 CCGTGTGTCAGAAATCAGATGATTTGATGAGAAAGCAAAAGCTTCAATGATGTT 720
DB 661 CCGTGTGTCAGAAATCAGATGATTTGATGAGAAAGCAAAAGCTTCAATGATGTT 720
QY 721 CTAGTGAATGTTAGTGGGATCTCCGCTCCGCCAATGCTATGCTTACATCATG 780
DB 721 CTAGTGAATGTTAGTGGGATCTCCGCTCCGCCAATGCTATGCTTACATCATG 780
QY 781 AAGAGATGAGATGTTAGATGAAAGCTTACAGTTTGTGAAGAAAGAAAGAAAGCTTACT 840
DB 781 AAGAGATGAGATGTTAGATGAAAGCTTACAGTTTGTGAAGAAAGAAAGAAAGCTTACT 840
QY 841 ATATCTCCAAATCTTCAATTTTCTGGGCCAACTCTGAGTATGAGAAAGATTAAGAAC 900
DB 841 ATATCTCCAAATCTTCAATTTTCTGGGCCAACTCTGAGTATGAGAAAGATTAAGAAC 900
QY 901 CAGACTGAGCATCAGAGGCCAAAGCAAACTCAAGCTGCTGACCTGAGAGCCAAAT 960
DB 901 CAGACTGAGCATCAGAGGCCAAAGCAAACTCAAGCTGCTGACCTGAGAGCCAAAT 960
QY 961 GAACTGTCTCTGCTCTCAGAGGGTGGACAGAAAGGAGAGCCCTCAGTCCACCC 1020
DB 961 GAACTGTCTCTGCTCTCAGAGGGTGGACAGAAAGGAGAGCCCTCAGTCCACCC 1020
QY 1021 TGTGCGGACTGTCTAAGCTCAGAGGAGAGCAAAAGGCCGCTGACCTCCGACGCTG 1080
DB 1021 TGTGCGGACTGTCTAAGCTCAGAGGAGAGCAAAAGGCCGCTGACCTCCGACGCTG 1080
QY 1081 CCCAGGCTGCCAGGCTGAGCGCTGCTTAAAGAGCAGGCCCTGTGACAGGCTGTC 1140
DB 1081 CCCAGGCTGCCAGGCTGAGCGCTGCTTAAAGAGCAGGCCCTGTGACAGGCTGTC 1140
QY 1141 AGTGGCTGACCTGTCCGAGACAGGCTGAGAGCAGCAATTAAGCTCAAGCTTCTTC 1200
DB 1141 AGTGGCTGACCTGTCCGAGACAGGCTGAGAGCAGCAATTAAGCTCAAGCTTCTTC 1200
QY 1201 TCTCTGATATCAATCAATCTTCAATTCAGCCAGCATGAGCAATCTTACATGCTTC 1260
DB 1201 TCTCTGATATCAATCAATCTTCAATTCAGCCAGCATGAGCAATCTTACATGCTTC 1260
QY 1261 TCTCTATCAAGAAATGCTTTGGAATTAATAAACTTCACTACTGATGAGCAAC 1320

DB 1261 TCTCTATCAAGAAATGCTTTGGAATTAATAAACTTCACTACTGATGAGCAAC 1320
QY 1321 AAGCTATGCCAGTTCTTCCCTGTTTCAAGAACTATGAGAGCAGATCCCGAAACAGTCT 1380
DB 1321 AAGCTATGCCAGTTCTTCCCTGTTTCAAGAACTATGAGAGCAGATCCCGAAACAGTCT 1380
QY 1381 GATAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGACAGCCGACGCTTCAAGAGCCAG 1440
DB 1381 GATAAGAGAAAGCCAGCATCTCCCAAGAAAGCTGACAGCCGACGCTTCAAGAGCCAG 1440
QY 1441 AGCAGCCATTCATCTGCTGAGAAACAGCAGCAGTGGCAACGAGAGTCTCTTTTA 1500
DB 1441 AGCAGCCATTCATCTGCTGAGAAACAGCAGCAGTGGCAACGAGAGTCTCTTTTA 1500
QY 1501 TCTCCAGTGCATGAGAGTGGGAGCGGTGAGAGCAATTAACACAGCTTCTTTTGGC 1560
DB 1501 TCTCCAGTGCATGAGAGTGGGAGCGGTGAGAGCAATTAACACAGCTTCTTTTGGC 1560
QY 1561 CTTTCACAGCAGCAGCAGCAGCAGCTTCAAGAGTCTGCTGCTGGGCTTAAAGGCTGGC 1620
DB 1561 CTTTCACAGCAGCAGCAGCAGCAGCTTCAAGAGTCTGCTGCTGGGCTTAAAGGCTGGC 1620
QY 1621 TCGGATATCTTGGGCCCCCAAGCTTACCTTCCCTGACAGCAGCTGGTATTTGGC 1680
DB 1621 TCGGATATCTTGGGCCCCCAAGCTTACCTTCCCTGACAGCAGCTGGTATTTGGC 1680
QY 1681 ACAGAGTCTCACTTCTGCTGCTGAGCTTACAGAGGAGAGGAGAGTACTTCT 1740
DB 1681 ACAGAGTCTCACTTCTGCTGCTGAGCTTACAGAGGAGAGGAGAGTACTTCT 1740
QY 1741 GCTTACAGCTCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
DB 1741 GCTTACAGCTCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
QY 1801 CAGAGGCAAGTGCAGAGGCTGAGTCTGCGGAGGAGCTGATGAGAGAGAGAGAGAGAG 1860
DB 1801 CAGAGGCAAGTGCAGAGGCTGAGTCTGCGGAGGAGCTGATGAGAGAGAGAGAGAGAG 1860
QY 1861 AAGCAGTTTAAACGACAGAGCTGCAATGAAATTTGAGAGAGCATCATGTAGAGAAC 1920
DB 1861 AAGCAGTTTAAACGACAGAGCTGCAATGAAATTTGAGAGAGCATCATGTAGAGAAC 1920
QY 1921 AGCTCAGGGAAGAGCTGCGGAAAGTGGGAGTCACTGATGCTTTTGGGAGCATGAA 1980
DB 1921 AGCTCAGGGAAGAGCTGCGGAAAGTGGGAGTCACTGATGCTTTTGGGAGCATGAA 1980
QY 1981 ATCATTGAGGCTCC 1995
DB 1981 ATCATTGAGGCTCC 1995

RESULT 2
US-10-377-072-27
Sequence 27, Application US/10377072
Publication No. US2004009501A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals Inc.
APPLICANT: Curtis, Rory A.J.
APPLICANT: Logan, Thomas Joseph
APPLICANT: Glucksmann, Maria A.
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Miyoung
APPLICANT: Teal, Feng-ying
TITLE OF INVENTION: NOVEL, 25869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
FILE REFERENCE: MP103-0180MNT
CURRENT APPLICATION NUMBER: US/10/377,072
CURRENT FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: US 09/895,860

QY 1681 ACAGAGTCTCACTTCTACTCTGCTCAGCCATCTTACGAGGAGTGGCACTTACTCT 1740
|
|
|
Db 1681 ACAGAGTCTCACTTCTACTCTGCTCAGCCATCTTACGAGGAGTGGCACTTACTCT 1740
|
|
|
QY 1741 GCTTACAGCTGAGCAGCTGCTGCCACTTTCGAGAGCAAGTCTATTCTGTGCGAGCGG 1800
|
|
|
Db 1741 GCTTACAGCTGAGCAGCTGCTGCCACTTTCGAGAGCAAGTCTATTCTGTGCGAGCGG 1800
|
|
|
QY 1801 CAGAGGCAAGTGAACAGAGCTGACTCGGCGGAGCTGGCATGAAAGAGCCCTTTTAA 1860
|
|
|
Db 1801 CAGAGGCAAGTGAACAGAGCTGACTCGGCGGAGCTGGCATGAAAGAGCCCTTTTAA 1860
|
|
|
QY 1861 AAGCAGTTTAAACGAGAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCAGAGAC 1920
|
|
|
Db 1861 AAGCAGTTTAAACGAGAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCAGAGAC 1920
|
|
|
QY 1921 AGCTACCGGAAAGCTGGGGGAAAGTGGGCACTGAGCTTTTTCGGGCAAGCATGAA 1980
|
|
|
Db 1921 AGCTACCGGAAAGCTGGGGGAAAGTGGGCACTGAGCTTTTTCGGGCAAGCATGAA 1980
|
|
|
QY 1981 ATCATTTGAGGTCTCC 1995
|
|
|
Db 1981 ATCATTTGAGGTCTCC 1995
|
|
|

RESULT 3

US-10-257-026-1
; Sequence 1, Application US/10257026
; Publication No. US20040086859A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10KDM5
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07
; NUMBER OF SEQ. ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ. ID NO 1
; LENGTH: 3059
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)..(2121)
US-10-257-026-1

Query Match 94.9%; Score 1893; DB 17; Length 3059;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGATGATGTTGAACTCAATTTGTTACTGAGAGGTTGGCTCTGCTGGAA 60
|
|
|
Db 127 ATGGCCCATGATGATGTTGAACTCAATTTGTTACTGAGAGGTTGGCTCTGCTGGAA 186
|
|
|
QY 61 AGTGAAGGGAAGAAAGTGTGCTTAATGATGATACCGGCCCAATTTGTGAATACAAATCATCC 120
|
|
|
Db 187 AGTGAAGGGAAGAAAGTGTGCTTAATGATGATACCGGCCCAATTTGTGAATACAAATCATCC 246
|
|
|
QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAAAGCTTATGAGCGAAAGTTGGAAG 180
|
|
|
Db 247 CACATTTTGAAGCCATTATATCACTGCTCCAAAGCTTATGAGCGAAAGTTGGAAG 306
|
|
|
QY 181 GACAAAGTTTATTAACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
|
|
|
Db 307 GACAAAGTTTATTAACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 366
|
|
|
QY 241 TGCAGTCAAGAGTTGTGTTAAGATGAAGAGTCCCAAGATGTGGCTCTCTCTTCA 300
|
|
|
Db 367 TGCAGTCAAGAGTTGTGTTAAGATGAAGAGTCCCAAGATGTGGCTCTCTCTTCA 426
|
|
|
QY 301 GACTGTTTCTCACTGTACTTCTGGGTAAAGCTGAGAGAGCTTCAACTCTGTCACTG 360
|
|
|

Db 427 GACTGTTTCTCACTGTACTTCTGGGTAAAGCTGAGAGAGCTTCAACTCTGTCACTG 486
|
|
|
QY 361 CTTCGAGTGGGTTTCTGAGTTCTCTGTTTCCCTGGGCTGTGTAAGAAATTC 420
|
|
|
Db 487 CTTCGAGTGGGTTTCTGAGTTCTCTGTTTCCCTGGGCTGTGTAAGAAATTC 546
|
|
|
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTCTTACCTGTTGCCAAATTTGGGCAAC 480
|
|
|
Db 547 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTCTTACCTGTTGCCAAATTTGGGCAAC 606
|
|
|
QY 481 CGAATTTCTCCAAATCTTTATCTTGCTGGCGGAGATGTCCTCAACAGAGAGCTGATA 540
|
|
|
Db 607 CGAATTTCTCCAAATCTTTATCTTGCTGGCGGAGATGTCCTCAACAGAGAGCTGATG 666
|
|
|
QY 541 CAGCAAAATGGGATTTGTTATGTTAAATGCAAGCTTAACTGTCCTCAAGCTGACTTT 600
|
|
|
Db 667 CAGCAAAATGGGATTTGTTATGTTAAATGCAAGCTTAACTGTCCTCAAGCTGACTTT 726
|
|
|
QY 601 ATCCCGAGTCTCATTTTCTGCTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
|
|
|
Db 727 ATCCCGAGTCTCATTTTCTGCTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 786
|
|
|
QY 661 CCGTGTGGAACAAATCAGTGAATTTTCAATGGAAGGAAAGCCTCCAAATGATGCTT 720
|
|
|
Db 787 CCGTGTGGAACAAATCAGTGAATTTTCAATGGAAGGAAAGCCTCCAAATGATGATGCTT 846
|
|
|
QY 721 CTAGTGCACTGTTTAACTGGGATCTCCGCTCCGCAACATCGCTATCGCTTACATCATG 780
|
|
|
Db 847 CTAGTGCACTGTTTAACTGGGATCTCCGCTCCGCAACATCGCTATCGCTTACATCATG 906
|
|
|
QY 781 AAGAGATGGAACATGCTCTTATGATGAAGCTTACAGATTTTGAAGAAAGAAAGCTTACT 840
|
|
|
Db 907 AAGAGATGGAACATGCTCTTATGATGAAGCTTACAGATTTTGAAGAAAGAAAGCTTACT 966
|
|
|
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGCACTATGAGAGAAATTAAGAAC 900
|
|
|
Db 967 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGCACTATGAGAGAAATTAAGAAC 1026
|
|
|
QY 901 CAGACTGGAAGATGAGGGCCCAAGAGCAAACTCAAGTGTCTGCACTGAGAGAACCAAT 960
|
|
|
Db 1027 CAGACTGGAAGATGAGGGCCCAAGAGCAAACTCAAGTGTCTGCACTGAGAGAACCAAT 1086
|
|
|
QY 961 GAACCTGTCCTGTGCTGCTGAGAGGTTGGAACAGAAAGCGAGCCCTCACTCCACCC 1020
|
|
|
Db 1087 GAACCTGTCCTGTGCTGCTGAGAGGTTGGAACAGAAAGCGAGCCCTCACTCCACCC 1146
|
|
|
QY 1021 TGTGCGCACTCTGTACTCTGAGGCAAGGCAAGAGCCCGTGCATCCCGCAGCGTG 1080
|
|
|
Db 1147 TGTGCGCACTCTGTACTCTGAGGCAAGGCAAGAGCCCGTGCATCCCGCAGCGTG 1206
|
|
|
QY 1081 CCCAGCGTCCAGCGCGTGCAGCCGTGCTGTTTGAAGACAGCCCGTGTGTAACAGCGCTTC 1140
|
|
|
Db 1207 CCCAGCGTCCAGCGCGTGCAGCCGTGCTGTTTGAAGACAGCCCGTGTGTAACAGCGCTTC 1266
|
|
|
QY 1141 AGTGGCTGCACTGTGCTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
|
|
|
Db 1267 AGTGGCTGCACTGTGCTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1326
|
|
|
QY 1201 TCTCTGATATCAAAATCAAGTTTATATTAACGAGCATGCGAGCATCTTAACTGAGCTTC 1260
|
|
|
Db 1327 TCTCTGATATCAAAATCAAGTTTATATTAACGAGCATGCGAGCATCTTAACTGAGCTTC 1386
|
|
|
QY 1261 TCTCTCATGAGAAATGCTTTGGAATCTTCAAACTTCACTCTGTGATGAGATGGACCAAC 1320
|
|
|
Db 1387 TCTCTCATGAGAAATGCTTTGGAATCTTCAAACTTCACTCTGTGATGAGATGGACCAAC 1446
|
|
|
QY 1321 AAGCTATGCCAGTTCTCCCTGTTTCAAGAACTATCGGACAGACTCCGAAACAGACTCT 1380
|
|
|
Db 1447 AAGCTATGCCAGTTCTCCCTGTTTCAAGAACTATCGGACAGACTCCGAAACAGACTCT 1506
|
|
|
QY 1381 GATTAAGAGAGAGCCAGCATTCCTCCAAAGAACTGACAGCCGCAAGGCTTTCAGACAGCCAG 1440
|
|
|
Db 1507 GATTAAGAGAGAGCCAGCATTCCTCCAAAGAACTGACAGCCGCAAGGCTTTCAGACAGCCAG 1566
|
|
|

QY 1441 AGAAGCGATTGGATTGGATGAGAACGACGAGGAGGACGCGCCAGAGGTCCTTTA 1500
| | | | |
Db 1567 AGAAGCGATTGGATTGGATGAGAACGACGAGGAGGACGCGCCAGAGGTCCTTTA 1626
| | | | |
QY 1501 TCTCCATCCTCATGAGAGTGGAGCGCTGAGAGCAATTAACACAGCTTCTTTTGGC 1560
| | | | |
Db 1627 TCTCCATCCTCATGAGAGTGGAGCGCTGAGAGCAATTAACACAGCTTCTTTTGGC 1686
| | | | |
QY 1561 CTTTCCACGAGCAGACGACCTCAGCAAGTCTGCTGGCCCTTAAAGGCTGGCAC 1620
| | | | |
Db 1687 CTTTCCACGAGCAGACGACCTCAGCAAGTCTGCTGGCCCTTAAAGGCTGGCAC 1746
| | | | |
QY 1621 TGGGATATCTGGCCCCCGCCAGACCTTACCCCTTCCCTGACGAGAGCTGTATTTTGGC 1680
| | | | |
Db 1747 TGGGATATCTGGCCCCCGCCAGACCTTACCCCTTCCCTGACGAGAGCTGTATTTTGGC 1806
| | | | |
QY 1681 ACAGAGTCTCTACACTTCTACTGCTGCTGACGACCTTACGAGAGGCTGCTGCTTACTCT 1740
| | | | |
Db 1807 ACAGAGTCTCTACACTTCTACTGCTGCTGACGACCTTACGAGAGGCTGCTGCTTACTCT 1866
| | | | |
QY 1741 GCCTACAGCTGACGACGAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGGCAGGCGG 1800
| | | | |
Db 1867 GCCTACAGCTGACGACGAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGGCAGGCGG 1926
| | | | |
QY 1801 CAGAGCCAGATGACAGAGCTGACTGCTGGCGGCGGAGCTGGCATGAAAGAGCCCTTTGAA 1860
| | | | |
Db 1927 CAGAGCCAGATGACAGAGCTGACTGCTGGCGGCGGAGCTGGCATGAAAGAGCCCTTTGAA 1986
| | | | |
QY 1861 AAGCATTTAAACGAGAGGCTGCCAAATGGAATTTGAGAGAGCATGTCAGAGAAC 1920
| | | | |
Db 1987 AAGCATTTAAACGAGAGGCTGCCAAATGGAATTTGAGAGAGCATGTCAGAGAAC 2046
| | | | |
QY 1921 AGGTACGAGGAGAGGCTGGGAGAAAGTGGGAGTCACTTACTTTTGGGCGAGCATGAA 1980
| | | | |
Db 2047 AGGTACGAGGAGAGGCTGGGAGAAAGTGGGAGTCACTTACTTTTGGGCGAGCATGAA 2106
| | | | |
QY 1981 ATCATTGAGTCTCC 1995
| | | | |
Db 2107 ATCATTGAGTCTCC 2121
| | | | |

RESULT 4
US-09-964-277-1
; Sequence 1, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-1

Query Match 94.9%; Score 1893; DB 9; Length 3496;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 121 CACATTTTGAAGCCATTATATCAATGCTCCCAAGCTTATGAGCGAAGTTGCAACAG 180
| | | | |
Db 682 CACATTTTGAAGCCATTATATCAATGCTCCCAAGCTTATGAGCGAAGTTGCAACAG 741
| | | | |
QY 181 GACAAAGTTAATTACAGAGCTCATTCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
| | | | |
Db 742 GACAAAGTTAATTACAGAGCTCATTCAGCATTCAGCGAAACATTAAGTTGACATTGAT 801
| | | | |
QY 241 TGCAGTCAGAAAGTTGATTTACATGAAACCTCCAAAGATGGTCCCTCTCTCTTCA 300
| | | | |
Db 802 TGCAGTCAGAAAGTTGATTTACATGAAACCTCCAAAGATGGTCCCTCTCTCTTCA 861
| | | | |
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGCTTCAACTCTGTTCACTG 360
| | | | |
Db 862 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGCTTCAACTCTGTTCACTG 921
| | | | |
QY 361 CTGGAGGTGGGTTTGTCTGATGTTCTCTGTTTCCCTGGCTCTGTGAAGAAATCC 420
| | | | |
Db 922 CTGGAGGTGGGTTTGTCTGATGTTCTCTGTTTCCCTGGCTCTGTGAAGAAATCC 981
| | | | |
QY 421 ACTCTAGTCCCTACCTGCACTTCTCAGCTTCTGCTTACCTGTTGCCAACTTGGGCCAAC 480
| | | | |
Db 982 ACTCTAGTCCCTACCTGCACTTCTCAGCTTCTGCTTACCTGTTGCCAACTTGGGCCAAC 1041
| | | | |
QY 481 CGAATTTCTTCCCAATCTTTATCTTGGCTGCGAGCGAGATGTCCTCAACAAAGAGCTGATA 540
| | | | |
Db 1042 CGAATTTCTTCCCAATCTTTATCTTGGCTGCGAGCGAGATGTCCTCAACAAAGAGCTGATA 1101
| | | | |
QY 541 CAGCAAGATGGAGTTGTTATGTTAAATGCGACGTATACCTGTCCAAAGCTGACTT 600
| | | | |
Db 1102 CAGCAAGATGGAGTTGTTATGTTAAATGCGACGTATACCTGTCCAAAGCTGACTT 1161
| | | | |
QY 601 ATCCCGAGTCTCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
| | | | |
Db 1162 ATCCCGAGTCTCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1221
| | | | |
QY 661 CCGTGGTTGGAACAATCAGTATGATTCATTGAGAAAGCAAAAGCTCCAAATGATGTT 720
| | | | |
Db 1222 CCGTGGTTGGAACAATCAGTATGATTCATTGAGAAAGCAAAAGCTCCAAATGATGTT 1281
| | | | |
QY 721 CTATGCACTGTTTATGCTGGGATCTCCGCTCCGCAACCATGCTATGCTTACATCATG 780
| | | | |
Db 1282 CTATGCACTGTTTATGCTGGGATCTCCGCTCCGCAACCATGCTATGCTTACATCATG 1341
| | | | |
QY 781 AAGAGATGACATGCTTTTATGATGAGCTTACAGATTTTGTAAAGAAAGAAAGAAAGCTTACT 840
| | | | |
Db 1342 AAGAGATGACATGCTTTTATGATGAGCTTACAGATTTTGTAAAGAAAGAAAGAAAGCTTACT 1401
| | | | |
QY 841 ATATCTCCAACTCAATTTTCTGGGCGCAACTCTGGACTATGAGAAAGATTAGAAC 900
| | | | |
Db 1402 ATATCTCCAACTCAATTTTCTGGGCGCAACTCTGGACTATGAGAAAGATTAGAAC 1461
| | | | |
QY 901 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTCAGCTGAGAGAGCCCAAT 960
| | | | |
Db 1462 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTCAGCTGAGAGAGCCCAAT 1521
| | | | |
QY 961 GAACTGTCTCTGCTGTCTCAGAGGTGAGCAAGAAAGGAGACGCGCTCAGTCCACCC 1020
| | | | |
Db 1522 GAACTGTCTCTGCTGTCTCAGAGGTGAGCAAGAAAGGAGACGCGCTCAGTCCACCC 1581
| | | | |
QY 1021 TGTGCGGACTGTGCTACCTCAGAGGAGAGCAAAAGGCGGTGATCCGCGACGCGT 1080
| | | | |
Db 1582 TGTGCGGACTGTGCTACCTCAGAGGAGAGCAAAAGGCGGTGATCCGCGACGCGT 1641
| | | | |
QY 1081 CCCAGGTCGCCAGGCTGAGCGCTGCTGTTAGAGAGCAGCCGCTGTGATCAGGCGCTC 1140
| | | | |
Db 1642 CCCAGGTCGCCAGGCTGAGCGCTGCTGTTAGAGAGCAGCCGCTGTGATCAGGCGCTC 1701
| | | | |
QY 1141 AGTGGCTGCACTGTCCGCAAGAGCTGAGAGACAGCAATTAAGTCAAGGCTTCTTC 1200
| | | | |
Db 1702 AGTGGCTGCACTGTCCGCAAGAGCTGAGAGACAGCAATTAAGTCAAGGCTTCTTC 1761
| | | | |
QY 1201 TCTCTGATATCAATCAAGTTTCATATTGAGCGAGATGAGCATCTTACATGGCTTC 1260
| | | | |

```

Db      1762  TCTCTGATATCAAAATCAGTTCTATATTCAGCCAGATGAGATCTTACATGGCTTC 1821
Qy      1261  TCCCATCAGAAAGTGTGGTGGAAATACTACAAACCTTCCACTACTCTGGATGGAGCAAC 1320
Db      1822  TCCCATCAGAAAGTGTGGTGGAAATACTACAAACCTTCCACTACTCTGGATGGAGCAAC 1881
Qy      1321  AAGCTATGCCAGTTCTCCCTGTTCAGAGACTATCGAGAGAGACTCCCGAAACAGTCTC 1380
Db      1882  AAGCTATGCCAGTTCTCCCTGTTCAGAGACTATCGAGAGAGACTCCCGAAACAGTCTC 1941
Qy      1381  GATTAAGAGAGAGAGAGAGAGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db      1942  GATTAAGAGAGAGAGAGAGAGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2001
Qy      1441  AGAAGAGAGATTTGATTTGGTTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db      2002  AGAAGAGAGATTTGATTTGGTTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2061
Qy      1501  TCTCACTGATCGAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db      2062  TCTCACTGATCGAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2121
Qy      1561  CTTTCCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db      2122  CTTTCCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2181
Qy      1621  TCGGATATCTTTGGCCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db      2182  TCGGATATCTTTGGCCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2241
Qy      1681  ACAGAGTCTCTACACTTCTACTCTGCTCAGCCATCAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db      2242  ACAGAGTCTCTACACTTCTACTCTGCTCAGCCATCAGAGAGAGAGAGAGAGAGAGAGAG 2301
Qy      1741  GCTTAAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db      2302  GCTTAAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2361
Qy      1801  CAGAACCCAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db      2362  CAGAACCCAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2421
Qy      1861  AAGCAGTTTAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db      2422  AAGCAGTTTAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2481
Qy      1921  AAGTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
Db      2482  AAGTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2541
Qy      1981  ATCATTGAGTCTCC 1995
Db      2542  ATCATTGAGTCTCC 2556

```

```

RESULT 5
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589) ... (2583)
US-09-816-494-1

Query Match      94.9%; Score 1893; DB 9; Length 3544;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  ATGGCCCATGAGATATTTGAACTCAAAATTTGTAAGAGAGAGTGGTCTGCTGGAA 60
Db      589  ATGGCCCATGAGATATTTGAACTCAAAATTTGTAAGAGAGTGGTCTGCTGGAA 648
Qy      61  AGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db      649  AGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 708
Qy      121  CACATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db      709  CACATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
Qy      181  GACAAAGTGTAAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db      769  GACAAAGTGTAAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
Qy      241  TGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db      829  TGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888
Qy      301  GACTGTTTCTACAGTCTCTGAGTAACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db      889  GACTGTTTCTACAGTCTCTGAGTAACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948
Qy      361  CTTGAGAGTGGAGTGTGAGAGTCTCTGTTTCCCTGAGCTCTGAGAGAGAGAGAGAGAGAG 420
Db      949  CTTGAGAGTGGAGTGTGAGAGTCTCTGTTTCCCTGAGCTCTGAGAGAGAGAGAGAGAGAG 1008
Qy      421  ACTCAAGTCTTACCTGATTTTCAAGCTTGTCTTCTGTTTCCCTGAGCTCTGAGAGAGAGAG 480
Db      1009  ACTCAAGTCTTACCTGATTTTCAAGCTTGTCTTCTGTTTCCCTGAGCTCTGAGAGAGAGAG 1068
Qy      481  CGAATCTTCCCAATCTTATCTTGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db      1069  CGAATCTTCCCAATCTTATCTTGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1128
Qy      541  CAGAGAAATGGAGATTTGTTATGTTAAATGCGAGTAACTGTCCAAAGCTGACTTT 600
Db      1129  CAGAGAAATGGAGATTTGTTATGTTAAATGCGAGTAACTGTCCAAAGCTGACTTT 1188
Qy      601  ATCCCGAGTCTCAATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db      1189  ATCCCGAGTCTCAATTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1248
Qy      661  CCGTGTGTTGACAATCAGTATGTTTATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db      1249  CCGTGTGTTGACAATCAGTATGTTTATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1308
Qy      721  CTAGTGAATGTTTAAAGTGGAGATCTCCGCTCGGCAATCGCTATCGCTATCATCATG 780
Db      1309  CTAGTGAATGTTTAAAGTGGAGATCTCCGCTCGGCAATCGCTATCGCTATCATCATG 1368
Qy      781  AAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db      1369  AAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1428
Qy      841  ATATCTCCAAATCTCAATTTTCTGGGCAACTCTGAGCTATGAGAGAGAGAGAGAGAGAGAG 900
Db      1429  ATATCTCCAAATCTCAATTTTCTGGGCAACTCTGAGCTATGAGAGAGAGAGAGAGAGAGAG 1488
Qy      901  CAGAGTGAAGATCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960

```


241 TGCAGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTGCTCTCTCTTGA 300
829 TGCAGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTGCTCTCTCTTGA 888
301 GACTGTTTTCTCACTGACTTCTGGGTAACTGGAGAGAGCTTCAACTCTGTTCACTG 360
889 GACTGTTTTCTCACTGACTTCTGGGTAACTGGAGAGAGCTTCAACTCTGTTCACTG 948
361 CTTCGAGGTGGGTTTCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
949 CTTCGAGGTGGGTTTCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 1008
421 ACTCTAGTCCCTACCTGCAATTTCTGAGCTTCTGAGCTTCACTGTTGCCAAATGGGCAAC 480
1009 ACTCTAGTCCCTACCTGCAATTTCTGAGCTTCTGAGCTTCACTGTTGCCAAATGGGCAAC 1068
481 CGAATTTCTTCCCAATCTTTATCTTGGCTGCCAGCAGATGTCTCTCAACAGAGCTGATA 540
1069 CGAATTTCTTCCCAATCTTTATCTTGGCTGCCAGCAGATGTCTCTCAACAGAGCTGATA 1128
541 CAGCAAAATGGGATTTGTTATGTGTTAAATGCACTAATCTGTCCAAAGCTGACTTT 600
1129 CAGCAAAATGGGATTTGTTATGTGTTAAATGCACTAATCTGTCCAAAGCTGACTTT 1188
601 ATCCCGAGTCTCAATTTCCGCGTGTGCTGTAATGACAGCTTTGTGAGAAATTTTG 660
1189 ATCCCGAGTCTCAATTTCCGCGTGTGCTGTAATGACAGCTTTGTGAGAAATTTTG 1248
661 CCGTGGTTGGAACAATCAGTAATTTCAATTGAGAAAGCAAAAGCTCCAAATGATGTGT 720
1249 CCGTGGTTGGAACAATCAGTAATTTCAATTGAGAAAGCAAAAGCTCCAAATGATGTGT 1308
721 CTAAGTCACTGTTTGTGCTGGGATCTCCGCTCCGCAACATGCTATGCTTCAATCAAT 780
1309 CTAAGTCACTGTTTGTGCTGGGATCTCCGCTCCGCAACATGCTATGCTTCAATCAAT 1368
781 AAGAGATGAGATGCTTTAGATGAGCTTACAGATTTGTGAGAAAGAAAGAAAGCTTACT 840
1369 AAGAGATGAGATGCTTTAGATGAGCTTACAGATTTGTGAGAAAGAAAGAAAGCTTACT 1428
841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTCTGACTATGAGAAAGATTAAAG 900
1429 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTCTGACTATGAGAAAGATTAAAG 1488
901 CAGACTGAGGATCAAGGGCCAAAGCAAACTCAAGCTGTGCTGAGAGGCAAT 960
1489 CAGACTGAGGATCAAGGGCCAAAGCAAACTCAAGCTGTGCTGAGAGGCAAT 1548
961 GAACCTGTCTCTGTCTCTCAGAGGTTGACAGAAAGGAGAGGCTCTCAGTCCACCC 1020
1549 GAACCTGTCTCTGTCTCTCAGAGGTTGACAGAAAGGAGAGGCTCTCAGTCCACCC 1608
1021 TGTGCGCACTGTCTCTCTCAGAGGCAAGCAAAAGGCTGTGATCCGCCAGCGTG 1080
1609 TGTGCGCACTGTCTCTCTCAGAGGCAAGCAAAAGGCTGTGATCCGCCAGCGTG 1668
1081 CCCAGCGTCCAGGCTGTGAGCGCTGTGTTAGAGAGAGCGCTGTGTAAGCGCTG 1140
1669 CCCAGCGTCCAGGCTGTGAGCGCTGTGTTAGAGAGAGCGCTGTGTAAGCGCTG 1728
1141 AGTGGGTCGACCTGTCCGACAGAGGCTGAGAAAGCAATTAAGCTTCAAGCTTCTTC 1200
1729 AGTGGGTCGACCTGTCCGACAGAGGCTGAGAAAGCAATTAAGCTTCAAGCTTCTTC 1788
1201 TCTCTGATATCAATCAGTTTCAATTTCAAGCCAGATGCGCATCTTTCATGAGCTTC 1260
1789 TCTCTGATATCAATCAGTTTCAATTTCAAGCCAGATGCGCATCTTTCATGAGCTTC 1848
1261 TCTCTATCAGAGAGCTTTTGAATTAATAAAGCTTCACTCACTGTGATGGAGCAAC 1320
1849 TCTCTATCAGAGAGCTTTTGAATTAATAAAGCTTCACTCACTGTGATGGAGCAAC 1908

QY 1321 AAGCTATGCAAGTTCTCCCTCTTTCAGGAATATCGAGAGAGACTCCGAAACAGACTCT 1380
Db 1909 AAGCTATGCAAGTTCTCCCTCTTTCAGGAATATCGAGAGAGACTCCGAAACAGACTCT 1968
QY 1381 GATTAAGAGGAAGCCAGCATCCCAAGAGCTGACAGCCGCGGCTTTACAGACAGCAG 1440
Db 1969 GATTAAGAGGAAGCCAGCATCCCAAGAGCTGACAGCCGCGGCTTTACAGACAGCAG 2028
QY 1441 AGCAAGCATTTGATTTGGGTGAGAACCGAGAGATGGGACCGCCAGAGGCTCTTTTA 1500
Db 2029 AGCAAGCATTTGATTTGGGTGAGAACCGAGAGATGGGACCGCCAGAGGCTCTTTTA 2088
QY 1501 TCTCCACTGCATCGAAGTGGAGCGTGGAGGCAATTAACACAGCAGCTTCTTTTGGC 1560
Db 2089 TCTCCACTGCATCGAAGTGGAGCGTGGAGGCAATTAACACAGCAGCTTCTTTTGGC 2148
QY 1561 CTTTCACAGCAGACAGACCTTCAGAGTCTGTGCTGAGCCTTGAAGGGCTGGCAG 1620
Db 2149 CTTTCACAGCAGACAGACCTTCAGAGTCTGTGCTGAGCCTTGAAGGGCTGGCAG 2208
QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGAGCTGTGATTTGCC 1680
Db 2209 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGAGCTGTGATTTGCC 2268
QY 1681 ACAGAGCTCTCACTTCTACTGCTGCTCAGCATCTACAGAGAGAGCTGCCAGTTACTCT 1740
Db 2269 ACAGAGCTCTCACTTCTACTGCTGCTCAGCATCTACAGAGAGAGCTGCCAGTTACTCT 2328
QY 1741 GCTTACAGCTGAGCAGCTGCTGCCAATTGGCGAGACCAAGTCTATTCTGTGCGCAGCGG 1800
Db 2329 GCTTACAGCTGAGCAGCTGCTGCCAATTGGCGAGACCAAGTCTATTCTGTGCGCAGCGG 2388
QY 1801 CAGAAAGCAAGTGCAGAGCTGACTCGCGGGGAGAGCTGGCAATGAAGAGCCCTTTGAA 1860
Db 2389 CAGAAAGCAAGTGCAGAGCTGACTCGCGGGGAGAGCTGGCAATGAAGAGAGCCCTTTGAA 2448
QY 1861 AAGCATTTAAACGAGAGCTGCCAAATGGAATTTGGAGAGAGCATGTCAGAGAAC 1920
Db 2449 AAGCATTTAAACGAGAGCTGCCAAATGGAATTTGGAGAGAGCATGTCAGAGAAC 2508
QY 1921 AGGTACCGGAAAGAGCTGGGAAAGTGGGAGTCACTTCTTTCGGGACGATGGAA 1980
Db 2509 AGGTACCGGAAAGAGCTGGGAAAGTGGGAGTCACTTCTTTCGGGACGATGGAA 2568
QY 1981 ATCATTGAGGTCTCC 1995
Db 2569 ATCATTGAGGTCTCC 2583

RESULT 7
US-10-425-114-26234
; Sequence 26234, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kowalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313) B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26234
; LENGTH: 3625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-028-H6_FLI
US-10-425-114-26234


```
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom
APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
APPLICANT: YAO, Monique G.; BURFORD, Neil
APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
APPLICANT: LEE, Ernestine A.; HAFALIA, April J.A.
APPLICANT: LU, Dying Alina M.; TRIBOULEY, Catherine M.
APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
APPLICANT: YUE, Henry; WARREN, Bridget A.
APPLICANT: NGUYEN, Daniel B.; CHAWLA, Narinder K.
APPLICANT: KEARNEY, Liam
TITLE OF INVENTION: PROTEIN PHOSPHATASES
FILE REFERENCE: PI-0173 PCT
CURRENT APPLICATION NUMBER: US/10/343,357
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: PCT/US01/23716
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,679
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/223,272
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/224,309
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 60/226,728
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 60/229,254
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: US 60/231,366
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 17
LENGTH: 3766
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CB1
US-10-343-357-17

Query Match      94.9%; Score 1893; DB 13; Length 3766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATGGCCCATGATGATGGAAGTCAATTTGTAAGAGAGTGGTGGCTGTGCGAA 60
DB      538 ATGGCCCATGATGATGGAAGTCAATTTGTAAGAGAGTGGTGGCTGTGCGAA 597
QY      61 AGTGAACGGAAAAAGTGTCTGTAATTGATAGCCGGCCATTTGTGAATACATCATCC 120
DB      598 AGTGAACGGAAAAAGTGTCTGTAATTGATAGCCGGCCATTTGTGAATACATCATCC 657
QY      121 CACATTTTGGAGCCCTTAATATCACTGCTCCAAAGCTTATGAAGGAAGTTGCAACAG 180
DB      658 CACATTTTGGAGCCCTTAATATCACTGCTCCAAAGCTTATGAAGGAAGTTGCAACAG 717
QY      181 GACAAAGTGTATTACAGAGCTCATCCGATTCAGCGAAGCATTAAGTTGACCTTAT 240
DB      718 GACAAAGTGTATTACAGAGCTCATCCGATTCAGCGAAGCATTAAGTTGACCTTAT 777
QY      241 TGCAGTCAGAAAGTGTATTAGATCAAGAGCTCCCAAGATGTTGCTCTCTTCA 300
DB      778 TGCAGTCAGAAAGTGTATTAGATCAAGAGCTCCCAAGATGTTGCTCTCTTCA 837
QY      301 GACTTTTTCACGTACTTCTGGGTAACTGGGAAGAGCTTTAACTCTGTTCACCTG 360
DB      838 GACTTTTTCACGTACTTCTGGGTAACTGGGAAGAGCTTTAACTCTGTTCACCTG 897
QY      361 CTTCGAGGTGGGTGTGCTGAGTTCTCTGTTGTTTCCCTGGGCTGTGTGAAGAAATCC 420
DB      898 CTTCGAGGTGGGTGTGCTGAGTTCTCTGTTGTTTCCCTGGGCTGTGTGAAGAAATCC 957
```

```
QY      421 ACTTATGTCCTTACCTGCAATTTCTCAGCCTTGCTTAACCTGTTGCCAATGGGCCAAC 480
DB      958 ACTTATGTCCTTACCTGCAATTTCTCAGCCTTGCTTAACCTGTTGCCAATGGGCCAAC 1017
QY      481 CGAATTTTCCCAATCTTTATCTGAGTCCAGCAGATGTCCTCAACAGAGGCTGATA 540
DB      1018 CGAATTTTCCCAATCTTTATCTGAGTCCAGCAGATGTCCTCAACAGAGGCTGATA 1077
QY      541 CAGCAAGATGGGATTGGTTATGTTAAATGCGACCTATACCTGTCCAAAGCTTACT 600
DB      1078 CAGCAAGATGGGATTGGTTATGTTAAATGCGACCAATACCTGTCCAAAGCTTACT 1137
QY      601 ATCCCCGAGTCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB      1138 ATCCCCGAGTCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
QY      661 CCGTGTGTGCAAAATCAGTAGATTTCTTTGAGAAAGAAAGCTCCCAATGATGTGT 720
DB      1198 CCGTGTGTGCAAAATCAGTAGATTTCTTTGAGAAAGAAAGCTCCCAATGATGTGT 1257
QY      721 CTATGTCATGTTTACCTGGGATCTCCGCTCCGCAACCATGCTATGCTTACATCATG 780
DB      1258 CTATGTCATGTTTACCTGGGATCTCCGCTCCGCAACCATGCTATGCTTACATCATG 1317
QY      781 AAGAGGATGGAATGCTTTTATGATGAACCTTAAGATTTGGAAGAAAGAAAGCTTACT 840
DB      1318 AAGAGGATGGAATGCTTTTATGATGAACCTTAAGATTTGGAAGAAAGAAAGCTTACT 1377
QY      841 ATATCTCCAACTTCAATTTTCTGGGCCCACTCTGCACTATGAGAAAGATTAAGAAC 900
DB      1378 ATATCTCCAACTTCAATTTTCTGGGCCCACTCTGCACTATGAGAAAGATTAAGAAC 1437
QY      901 CAGACTGAGATCAGGGGCCAAAGCAATCTCAAGTGTCTGCACTGGAGAGCCAAAT 960
DB      1438 CAGACTGAGATCAGGGGCCAAAGCAATCTCAAGTGTCTGCACTGGAGAGCCAAAT 1497
QY      961 GAACCTGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB      1498 GAACCTGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1557
QY      1021 TGTGCGCATCTGCTTACCTCAGAGGCAAGCAAAAGCCCGTGCATCCCGCCAGCGTG 1080
DB      1558 TGTGCGCATCTGCTTACCTCAGAGGCAAGCAAAAGCCCGTGCATCCCGCCAGCGTG 1617
QY      1081 CCCAGCGTCCAGCCGTCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB      1618 CCCAGCGTCCAGCCGTCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1677
QY      1141 AGTGGGCTGCACTGTCCGCAAGAGCTGGAAGCAAGCAATAGCTCAAGCGTTCCTTC 1200
DB      1678 AGTGGGCTGCACTGTCCGCAAGAGCTGGAAGCAAGCAATAGCTCAAGCGTTCCTTC 1737
QY      1201 TCTCTGATATCAAAATCAGTTTCAATATTCAGCCAGCATGAGCATCTTACATGGCTTC 1260
DB      1738 TCTCTGATATCAAAATCAGTTTCAATATTCAGCCAGCATGAGCATCTTACATGGCTTC 1797
QY      1261 TCTCTCATCAAGAAAGTGTGTTGGAATATCAAACTTTCATCTCTGATGAGGACCAAC 1320
DB      1798 TCTCTCATCAAGAAAGTGTGTTGGAATATCAAACTTTCATCTCTGATGAGGACCAAC 1857
QY      1321 AAGCTATGCAATGTTCTCCCTGTTTCAAGAACTATCGAGAGAGACTCCCGAAACAGTCT 1380
DB      1858 AAGCTATGCAATGTTCTCCCTGTTTCAAGAACTATCGAGAGAGACTCCCGAAACAGTCT 1917
QY      1381 GATTAAGAGAAAGCAGCATCTCCCAAGAGCTGCAAGCCGACGAGCTTTCAGACAGCAG 1440
DB      1918 GATTAAGAGAAAGCAGCATCTCCCAAGAGCTGCAAGCCGACGAGCTTTCAGACAGCAG 1977
QY      1441 AGCAAGGATTTGATTCGGTCAAGCAAGCAAGAGAGTGGCAACGAGGCTCCCTTTTA 1500
DB      1978 AGCAAGGATTTGATTCGGTCAAGCAAGCAAGAGAGTGGCAACGAGGCTCCCTTTTA 2037
QY      1501 TCTTCACTGCATGGAAGTGGAGCTGGAGGACAAATTACACACAGCTTCTTTTCGGC 1560
```

Db 2038 TCTCCACTGATGAGTGGGAGCGTGGAGCAATTACACACAGCTTCTTTTCGG 2097
QY 1561 CTTTCCACCAAGCCAGCAGCACTTCAAGAACTCTGTGGCTTAAAGGCTGAC 1620
Db 2098 CTTTCCACCAAGCCAGCAGCACTTCAAGAACTCTGTGGCTTAAAGGCTGAC 2157
QY 1621 TCGGATATCTTGGCCCCCGAGCCTTAACCCCTTCCCTGACGAGCCTGTATTTTGGC 1680
Db 2158 TCGGATATCTTGGCCCCCGAGCCTTAACCCCTTCCCTGACGAGCCTGTATTTTGGC 2217
QY 1681 ACAGATCTCTACACTTCTACTCTGCTCAGCCATCTACGAGAGGAGTGGCACTTACTCT 1740
Db 2218 ACAGATCTCTACACTTCTACTCTGCTCAGCCATCTACGAGAGGAGTGGCACTTACTCT 2277
QY 1741 GCTTACAGCTGACAGCCAGCTGCCCACTTGGGAGAACAGTATTTCTGTGGCAGGCG 1800
Db 2278 GCTTACAGCTGACAGCCAGCTGCCCACTTGGGAGAACAGTATTTCTGTGGCAGGCG 2337
QY 1801 CAGAAAGCAGTGAAGAGTGACTGCGCGGAGAGCTGGGATGAAGAGGCCCTTTGAA 1860
Db 2338 CAGAAAGCAGTGAAGAGTGACTGCGCGGAGAGCTGGGATGAAGAGGCCCTTTGAA 2397
QY 1861 AACGAGTTTAAAGCAGAAAGCTGCCAAATGGAATTTGAGAGAGCATCATGTAGAGAAC 1920
Db 2398 AACGAGTTTAAAGCAGAAAGCTGCCAAATGGAATTTGAGAGAGCATCATGTAGAGAAC 2457
QY 1921 AGGTCAAGGAGAGAGCTGGGAAAAGTGGGAGTCACTGCTTTTGGGAGCATGAA 1980
Db 2458 AGGTCAAGGAGAGAGCTGGGAAAAGTGGGAGTCACTGCTTTTGGGAGCATGAA 2517
QY 1981 ATGATTGAGGTCTCC 1995
Db 2518 ATGATTGAGGTCTCC 2532

RESULT 9

US-10-648-593-115
; Sequence 115, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 115
; LENGTH: 4790
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-648-593-115

Query Match 94.9%; Score 1893; DB 17; Length 4790;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGATGATGGAACCTCAATTTGTAAGAGGTTGTGCTCTGTGAA 60
Db 184 ATGGCCCATGATGATGGAACCTCAATTTGTAAGAGGTTGTGCTCTGTGAA 243
QY 61 AGTGGAAAGGAAAGTGTCTTAATGATAGCCGGCATTTTGAATCAATATACATCC 120
Db 244 AGTGGAAAGGAAAGTGTCTTAATGATAGCCGGCATTTTGAATCAATATACATCC 303
QY 121 CACATTTTGAAGCATTATATCAATGCTCCAGCTTATGAAGCAAGTTGCAACAG 180
Db 304 CACATTTTGAAGCATTATATCAATGCTCCAGCTTATGAAGCAAGTTGCAACAG 363

QY 181 GCAAAAGTTAATTACAGAGCTCATCCAGCATTTGCGGAAACATATAGTTGACATTGAT 240
Db 364 GCAAAAGTTAATTACAGAGCTCATCCAGCATTTGCGGAAACATATAGTTGACATTGAT 423
QY 241 TGCAGTCAGAAAGTTGTATTTACATCAATCAAGCTCCCAAGATGTTGCTCTCTTCA 300
Db 424 TGCAGTCAGAAAGTTGTATTTACATCAATCAAGCTCCCAAGATGTTGCTCTCTTCA 483
QY 301 GACTGTTTCTCACTGTACTTCTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACCTG 360
Db 484 GACTGTTTCTCACTGTACTTCTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACCTG 543
QY 361 CTTGCAAGTGGGTTTGTCTAGTTCTCTGTTGTTCCCGGCTCTGTGAAGGAAATCC 420
Db 544 CTTGCAAGTGGGTTTGTCTAGTTCTCTGTTGTTCCCGGCTCTGTGAAGGAAATCC 603
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACTGTTGCCAAATTTGGCCAAAC 480
Db 604 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACTGTTGCCAAATTTGGCCAAAC 663
QY 481 CGAATTTCTCCAAATCTTATCTTGGCTGCCAGGATGCTCTCAACAAAGAGCTGAT 540
Db 664 CGAATTTCTCCAAATCTTATCTTGGCTGCCAGGATGCTCTCAACAAAGAGCTGAT 723
QY 541 CAGCAAAATGGATTTGGTTATGTTAAATGCAAGTATACCTGTCCAAAGCCTGACTTT 600
Db 724 CAGCAAAATGGATTTGGTTATGTTAAATGCAAGTATACCTGTCCAAAGCCTGACTTT 783
QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Db 784 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 843
QY 661 CCGTGGTTGACAAATCAGTAAATTTCAATTGAGAAACAAAGCCTTCAATGATGTGT 720
Db 844 CCGTGGTTGACAAATCAGTAAATTTCAATTGAGAAACAAAGCCTTCAATGATGTGT 903
QY 721 CTAGTGCACTGTTTACGTGGATCTCCGCTCCGCCACATGCTATCCCTACATCATG 780
Db 904 CTAGTGCACTGTTTACGTGGATCTCCGCTCCGCCACATGCTATCCCTACATCATG 963
QY 781 AAGAGATGACATGCTTTAGTGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTACT 840
Db 964 AAGAGATGACATGCTTTAGTGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTACT 1023
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGACTATGAGAAAGATTTAAGAAC 900
Db 1024 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGACTATGAGAAAGATTTAAGAAC 1083
QY 901 CAGACTGAGCATCAGGCGCCAAAGCAAACTCAACTGTCTGCACTTGAGAGAACCCAAAT 960
Db 1084 CAGACTGAGCATCAGGCGCCAAAGCAAACTCAACTGTCTGCACTTGAGAGAACCCAAAT 1143
QY 961 GAACTGTCTCCGTCTCTCAGAGGTTGACAGAAAGGAGAGGCGCCCTCACTCCACCC 1020
Db 1144 GAACTGTCTCCGTCTCTCAGAGGTTGACAGAAAGGAGAGGCGCCCTCACTCCACCC 1203
QY 1021 TGTGCCGACTTGTCTCACTCAGAGGAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1080
Db 1204 TGTGCCGACTTGTCTCACTCAGAGGAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1263
QY 1081 CCCAGCTGCGCCAGCGTGCAGCGTGTGTTTGAAGACAGCCCGCTGTGATCAGGGGCTC 1140
Db 1264 CCCAGCTGCGCCAGCGTGCAGCGTGTGTTTGAAGACAGCCCGCTGTGATCAGGGGCTC 1323
QY 1141 AGTGGCTGCACTGTCCGAGAGCAGGCTGGAAGCAGCAATTAAGCTCAAGCTTCTTC 1200
Db 1324 AGTGGCTGCACTGTCCGAGAGCAGGCTGGAAGCAGCAATTAAGCTCAAGCTTCTTC 1383
QY 1201 TCTCTGATATCAAAATCAGTTTCAATTTAGCCAGATGAGCAGATCTTACATGCTTC 1260
Db 1384 TCTCTGATATCAAAATCAGTTTCAATTTAGCCAGATGAGCAGATCTTACATGCTTC 1443

QY 1261 TCCTCATCAGAAAGTCTTTGAAATACTACAAACCTTTCACCTACTGTGATGGGACCAAC 1320
1444 TCCTCATCAGAAAGTCTTTGAAATACTACAAACCTTTCACCTACTGTGATGGGACCAAC 1503
QY 1321 AAGCTATGCCAGTCTTCCCTGTTTCAGAACTATCGGAGCAGACTCCGGAACCACTGCT 1380
1504 AACCTATGCCAGTCTTCCCTGTTTCAGAACTATCGGAGCAGACTCCGGAACCACTGCT 1563
QY 1381 GATTAAGGAGGAGCAGCATCCCAAGAACTGACAGCCGCGCTTACAGACCCAG 1440
1564 GATTAAGGAGGAGCAGCATCCCAAGAACTGACAGCCGCGCTTACAGACCCAG 1623
QY 1441 AGCAAGGATTCGATTCGATCAGAACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1500
1624 AGCAAGGATTCGATTCGATCAGAACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1683
QY 1501 TCTTCACCTGCATGAAAGTGGAGCGTGGAGAGCAATTACACACAGCTTCTTTTCGCG 1560
1684 TCTTCACCTGCATGAAAGTGGAGCGTGGAGAGCAATTACACACAGCTTCTTTTCGCG 1743
QY 1561 CTTTCCACCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
1744 CTTTCCACCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1803
QY 1621 TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGATATTTTGGC 1680
1804 TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGATATTTTGGC 1863
QY 1681 ACAGAGTCTCTACACCTTCTACTCTGCTCAGCCATCTACGAGGAGCAGTGCAGTTACTCT 1740
1864 ACAGAGTCTCTACACCTTCTACTCTGCTCAGCCATCTACGAGGAGCAGTGCAGTTACTCT 1923
QY 1741 GCTTACAGCTGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1800
1924 GCTTACAGCTGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1983
QY 1801 CAGAACCCAGTACAGACGCTGCTCGCGCGAGCTGGCATGAAGAGCCCTTTGAA 1860
1984 CAGAACCCAGTACAGACGCTGCTCGCGCGAGCTGGCATGAAGAGCCCTTTGAA 2043
QY 1861 AAGCAGTTTAAACGAGAACTGCGCAATGGAATTTTGAAGAGCATCATGTCAGAGAAC 1920
2044 AAGCAGTTTAAACGAGAACTGCGCAATGGAATTTTGAAGAGCATCATGTCAGAGAAC 2103
QY 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTGATTTTCCGGCAGCATGGA 1980
2104 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTGATTTTCCGGCAGCATGGA 2163
QY 1981 ATCATTGAGGTCTCC 1995
2164 ATCATTGAGGTCTCC 2178
Db

RESULT 10
US-10-296-115-520
; Sequence 520, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 520
; LENGTH: 2966
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

NAME/KEY: misc.feature
LOCATION: (1)-(2966)
OTHER INFORMATION: n = a,t,c or g
US-10-296-115-520

Query Match 92.7%; Score 1849; DB 13; Length 2966;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1949; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 45 GGTGCTCTGCTGGAAGTGGAGCGAAAGTCTGCTAATTGATAGCGGCCATTGT 104
67 GGTGCTCTGCTGGAAGTGGAGCGAAAGTCTGCTAATTGATAGCGGCCATTGT 126
Db

QY 105 GGAATACATATATCCCATTTTGGAGCCATTAATATCACTGCTCCAGCTATGAA 164
127 GGAATACATATATCCCATTTTGGAGCCATTAATATCACTGCTCCAGCTATGAA 186
Db

QY 165 GCGAAGTTGCAACAGACCAAGTGTAAATTAACAGAGCTCATCCAGCAATTCAGCGAAACA 224
187 GCGAAGTTGCAACAGACCAAGTGTAAATTAACAGAGCTCATCCAGCAATTCAGCGAAACA 246
Db

QY 225 TAAGTTGACATTTGATGAGTCAAGAGTGTAGTTACGATCAAAGCTCCAAAGATGT 284
247 TAAGTTGACATTTGATGAGTCAAGAGTGTAGTTACGATCAAAGCTCCAAAGATGT 306
Db

QY 285 TGGCTCTCTCTGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344
307 TGGCTCTCTCTGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
Db

QY 345 CAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404
367 CAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
Db

QY 405 CTGTGAAGAAATATCACTCTATGCTCTAAGCTGATTTCTGAGCTTGTCTTACCTGTTGC 464
427 CTGTGAAGAAATATCACTCTATGCTCTAAGCTGATTTCTGAGCTTGTCTTACCTGTTGC 486
Db

QY 465 CAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 524
487 CAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546
Db

QY 525 CAACAGAGAGCTGATACAGCAGAAATGGATTTGTTATGTTAAATGACAGTATACCTG 584
547 CAACAGAGAGCTGATACAGCAGAAATGGATTTGTTATGTTAAATGACAGTATACCTG 606
Db

QY 585 TCCAAAGCTGATCTTTATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644
607 TCCAAAGCTGATCTTTATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666
Db

QY 645 TTTGAGGAAATTTTGGCGTGTGGAGCAATTCAGTATGTTCAATGGAAGAAAGCAAGC 704
667 TTTGAGGAAATTTTGGCGTGTGGAGCAATTCAGTATGTTCAATGGAAGAAAGCAAGC 726
Db

QY 705 CTCGAATGAGTGTCTAGTGAAGTGTGTTAGTGGAGTCTCCGCTCCGACCATGCTG 764
727 CTCGAATGAGTGTCTAGTGAAGTGTGTTAGTGGAGTCTCCGCTCCGACCATGCTG 786
Db

QY 765 TATGCTTACATCATGAAGAGATGAGACATGCTTTAGATGAAGCTTACAGATTTGTGAA 824
787 TATGCTTACATCATGAAGAGATGAGACATGCTTTAGATGAAGCTTACAGATTTGTGAA 846
Db

QY 825 AGAAAAAGACCTCTATATCTCCAAACTTAAATTTTCTGGGCAACCTCCGAGCTATGA 884
847 AGAAAAAGACCTCTATATCTCCAAACTTAAATTTTCTGGGCAACCTCCGAGCTATGA 906
Db

QY 885 GAAAGAGTTAAGAACAGACTGAGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGCA 944
907 GAAAGAGTTAAGAACAGACTGAGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGCA 966
Db

QY 945 CTGAGAGAGCCAAATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004
967 CTGAGAGAGCCAAATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
Db

QY 1005 GCCCTCAGTCCACCACTGTGCTCCACTCTGTACTGACCTAGAGGACGACGACCAAGGCCCGT 1064
DB 1027 GCGCTTCAGTCCACCACTGTGCTCCACTCTGTACTGACCTAGAGGACGACGACCAAGGCCCGT 1086
QY 1065 GCATCCCGGACGAGCGTCCGACGAGCGTCCGACGAGCGTCCGCTGTGTAGAGGACGACGCC 1124
DB 1087 GCATCCCGGACGAGCGTCCGACGAGCGTCCGACGAGCGTCCGCTGTGTAGAGGACGACGCC 1146
QY 1125 GCTGTGACAGGCGCTCAGTGGGCTGACCTGTGCTCCGACGACGAGCTGGAGACGACCAATA 1184
DB 1147 GCTGTGACAGGCGCTCAGTGGGCTGACCTGTGCTCCGACGACGAGCTGGAGACGACCAATA 1206
QY 1185 GCTCAAGCGTTCCTCTCTCTGTGATATCAATCAGTTTCAATTCAGCCGACGATGGCAGC 1244
DB 1207 GCTCAAGCGTTCCTCTCTCTGTGATATCAATCAGTTTCAATTCAGCCGACGATGGCAGC 1266
QY 1245 ATCTTACATAGGCTTCCTCTCTCATCAGAAAGATGCTTTGGAAATCTAACAACTTCCACTAC 1304
DB 1267 ATCTTACATAGGCTTCCTCTCTCATCAGAAAGATGCTTTGGAAATCTAACAACTTCCACTAC 1326
QY 1305 TCTGTGATGGGACCAACAGCTATGCGAGTTCTCCCTGTTCAGAAACTATCGAGACGAC 1364
DB 1327 TCTGTGATGGGACCAACAGCTATGCGAGTTCTCCCTGTTCAGAAACTATCGAGACGAC 1386
QY 1365 TCCCGAAACCAAGTCTGTATAGAGGAGGAGGACGACATCCCAAGAGCTGACGACGCGCAG 1424
DB 1387 TCCCGAAACCAAGTCTGTATAGAGGAGGAGGACGACATCCCAAGAGCTGACGACGCGCAG 1446
QY 1425 GCTTTCAGACAGCCAGACGACGAGGATTCATTCGCTCAGAACGACGACGAGTGGCAGC 1484
DB 1447 GCTTTCAGACAGCCAGACGAGGATTCATTCGCTCAGAACGACGAGTGGCAGC 1506
QY 1485 CCGAGAGTCCCTTTTATCTCCTCATCTGATCGAAGTGGAGGCTGAGAGCAATTAACACAC 1544
DB 1507 CCGAGAGTCCCTTTTATCTCCTCATCTGATCGAAGTGGAGGCTGAGAGCAATTAACACAC 1566
QY 1545 CAGCTTCCTTTTGGGCTTTCCACAGCCAGACGACCTCAGCAAGTCTGAGCGCTGGG 1604
DB 1567 CAGCTTCCTTTTGGGCTTTCCACAGCCAGACGACCTCAGCAAGTCTGAGCGCTGGG 1626
QY 1605 CCTTAAAGGCTGACCTCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACGAG 1664
DB 1627 CCTTAAAGGCTGACCTCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACGAG 1686
QY 1665 CAGCTGTATTTTGGCCACAGAGTCTCACTTCTACTCTGCTCAGCCACTTACCGAGG 1724
DB 1687 CAGCTGTATTTTGGCCACAGAGTCTCACTTCTACTCTGCTCAGCCACTTACCGAGG 1746
QY 1725 CAGTGCAGTTACTCTGCTCTACAGCTGACGACGCTGCCCACTTGGGAGAACCAAGCTTA 1784
DB 1747 CAGTGCAGTTACTCTGCTCTACAGCTGACGACGCTGCCCACTTGGGAGAACCAAGCTTA 1806
QY 1785 TTTCTGTGCGAGGCGGACGAGCCAAAGTGAACAGAGTGAAGTCTCGGCGGAGCTGGAGTGA 1844
DB 1807 TTTCTGTGCGAGGCGGACGAGCCAAAGTGAACAGAGTGAAGTCTCGGCGGAGCTGGAGTGA 1866
QY 1845 AGAGAGCCCTTTGAAAGAGAGTTTAAAGCAGAGAGTGCACAAATTTGAGAGAG 1904
DB 1867 AGAGAGCCCTTTGAAAGAGAGTTTAAAGCAGAGAGTGCACAAATTTGAGAGAG 1926
QY 1905 CATCATGTCAAGAGAGAGTCAAGGAGAGAGTGGGAGAGTGGGAGAGTGAAGTCTT 1964
DB 1927 CATCATGTCAAGAGAGAGTCAAGGAGAGAGTGGGAGAGTGGGAGAGTGAAGTCTT 1986
QY 1965 TTGCGGACGACGATGAGAAATCATTTGAGTCTCC 1995
DB 1987 TTGCGGACGACGATGAGAAATCATTTGAGTCTCC 2017

RESULT 11
US-10-168-506-2
; Sequence 2, Application US/10168506
; Publication No. US20040053229A1

; GENERAL INFORMATION:
; APPLICANT: PLOMMAN, GREGORY D.
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDASANAM, SUCHA
; APPLICANT: HILL, RON
; APPLICANT: FLANAGAN, PETER
; TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
; FILE REFERENCE: 038602/1351
; CURRENT APPLICATION NUMBER: US/10/168,506
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: PCT/US00/34736
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-168-506-2
Query Match 92.3%; Score 1842; DB 13; Length 2732;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATTTGGAACCTCAATTTGTTACTGAGAGGTTGGCTCTGCTGGAA 60
DB 538 ATGGCCCATGAGATGATTTGGAACCTCAATTTGTTACTGAGAGGTTGGCTCTGCTGGAA 597
QY 61 AGTGAACGGAAAGAGTGTGCTAATTTGATGAGCGGSCATTTTGTGGAATACATATCC 120
DB 598 AGTGAACGGAAAGAGTGTGCTAATTTGATGAGCGGSCATTTTGTGGAATACATATCC 657
QY 121 CACATTTTGAAGCCATTATATCAACTGCTCCAGCTTATGAGCGGAAGTTGCAACAG 180
DB 658 CACATTTTGAAGCCATTATATCAACTGCTCCAGCTTATGAGCGGAAGTTGCAACAG 717
QY 181 GACAAAGTGTATTTACAGAGCTCATCCGACATTCAGCGGAAACATTAAGTTGATGAT 240
DB 718 GACAAAGTGTATTTACAGAGCTCATCCGACATTCAGCGGAAACATTAAGTTGATGAT 777
QY 241 TGCAGTCAGAAAGTTTATGATTAAGATCAAGAGTCCCAAGATGTTGCTCTCTCTTA 300
DB 778 TGCAGTCAGAAAGTTTATGATTAAGATCAAGAGTCCCAAGATGTTGCTCTCTCTTA 837
QY 301 GACTGTTTCTGACTGATCTTCTGAGTAACTGAGAGAGCTTCACTCTGTTCACTG 360
DB 838 GACTGTTTCTGACTGATCTTCTGAGTAACTGAGAGAGCTTCACTCTGTTCACTG 897
QY 361 CTTCAGAGTGGGTTGCTGAGATTTCTGTTGTTTCCCTGGGCTCTGTGAGGAAATGC 420
DB 898 CTTCAGAGTGGGTTGCTGAGATTTCTGTTGTTTCCCTGGGCTCTGTGAGGAAATGC 957
QY 421 ACTCTAGTCCCTAACCGGATTTCTGAGCTTCTGAGCTTCACTGTTGCCAAATTTGGGCCAAC 480
DB 958 ACTCTAGTCCCTAACCGGATTTCTGAGCTTCTGAGCTTCACTGTTGCCAAATTTGGGCCAAC 1017
QY 481 CGAATTTCTCCCATCTTTATCTTGGCTCCAGCGAGATGCTCTCAACAAGAGCTGATA 540
DB 1018 CGAATTTCTCCCATCTTTATCTTGGCTCCAGCGAGATGCTCTCAACAAGAGCTGATA 1077
QY 541 CAGCAGAAATGGGATTTGTTATGTTTAAATGCCAGTATACCTGTGCCAAAGCTGACTT 600
DB 1078 CAGCAGAAATGGGATTTGTTATGTTTAAATGCCAGTATACCTGTGCCAAAGCTGACTT 1137
QY 601 ATCCCGAGTCTCATTTCCGCGTGCTGCTGAGATGACAGCTTTTGTGAGAAATTTTG 660
DB 1138 ATCCCGAGTCTCATTTCCGCGTGCTGCTGAGATGACAGCTTTTGTGAGAAATTTTG 1197
QY 661 CCGTGTGTGACCAATCAGTAGATTTCAATGAGAAAGCAAAAGCTTCAATGAGATGTT 720
DB 1198 CCGTGTGTGACCAATCAGTAGATTTCAATGAGAAAGCAAAAGCTTCAATGAGATGTT 1257

```
QY 721 CTATGCACTGTTTAAAGTGGATCTCCGCTCCGCCACATCGCTATCGGCTACATCATG 780
Db 1258 CTATGCACTGTTTAAAGTGGATCTCCGCTCCGCCACATCGCTATCGGCTACATCATG 1317
QY 781 AAGAGATGGAACATGCTTTTAAAGTGAAGCTTAAAGATTTTGAAGAAAAAGAAAGCTTACT 840
Db 1318 AAGAGATGGAACATGCTTTTAAAGTGAAGCTTAAAGATTTTGAAGAAAAAGAAAGCTTACT 1377
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGCACTATGAAAGAAATTAAGAAC 900
Db 1378 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGCACTATGAAAGAAATTAAGAAC 1437
QY 901 CAGACTGAGCATCAGGGCCAAAGAGCAATCAAGCTGTGACCTGAGAGCAACCAAT 960
Db 1438 CAGACTGAGCATCAGGGCCAAAGAGCAATCAAGCTGTGACCTGAGAGCAACCAAT 1497
QY 961 GAACTGTCTCTGTCTCTAGAGGGTGAACAGAAAAAGCGAGACGCTCCCTCACTCCACC 1020
Db 1498 GAACTGTCTCTGTCTCTAGAGGGTGAACAGAAAAAGCGAGACGCTCCCTCACTCCACC 1557
QY 1021 TGTGCCGACTCTCTCTCTCTAGAGGGTGAACAGAAAAAGCGAGACGCTCCCTCACTCC 1080
Db 1558 TGTGCCGACTCTCTCTCTCTAGAGGGTGAACAGAAAAAGCGAGACGCTCCCTCACTCC 1617
QY 1081 CCCAGGCTGCGCCAGCGCTGACGCGCTGTAGAGGACAGCGCGGTGATCAGAGGCGTC 1140
Db 1618 CCCAGGCTGCGCCAGCGCTGACGCGCTGTAGAGGACAGCGCGGTGATCAGAGGCGTC 1677
QY 1141 AGTGGCTGTGACCTGTCTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCGTTCTTC 1200
Db 1678 AGTGGCTGTGACCTGTCTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCGTTCTTC 1737
QY 1201 TCTCTGATATCAAAATCAATTTTCAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1260
Db 1738 TCTCTGATATCAAAATCAATTTTCAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1797
QY 1261 TCTCTGATATCAAAATCAATTTTCAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1320
Db 1798 TCTCTGATATCAAAATCAATTTTCAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1857
QY 1321 AAGCTATGCGAGTTTCTCCCTGTTTCAAGAACTATCGAGACAGCTCCGAAACCACTGCT 1380
Db 1858 AAGCTATGCGAGTTTCTCCCTGTTTCAAGAACTATCGAGACAGCTCCGAAACCACTGCT 1917
QY 1381 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1918 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1977
QY 1441 AGCAAGCATTGATTCGATTCAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1978 AGCAAGCATTGATTCGATTCAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2037
QY 1501 TCTCCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 2038 TCTCCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2097
QY 1561 CTTTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 2098 CTTTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2157
QY 1621 TCGATATCTTGGCCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 2158 TCGATATCTTGGCCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2217
QY 1681 ACAGAGTCTCTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Db 2218 ACAGAGTCTCTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2277
QY 1741 GGTCTACGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 2278 GGTCTACGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2337
```

```
QY 1801 CAGAAAGCCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db 2338 CAGAAAGCCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2397
QY 1861 AAGCAGTTTAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 2398 AAGCAGTTTAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2457
QY 1921 AGGTGACGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
Db 2458 AGGTGACGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2517
QY 1981 ATCATTGAGGTCTCC 1995
Db 2518 ATCATTGAGGTCTCC 2532

RESULT 12
US-10-094-749-673
; Sequence 673, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOMYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 08435/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 673
; LENGTH: 2102
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-673

Query Match 89.8%; Score 1791; DB 16; Length 2102;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAGGCCCATGAGATGAGTGAACCAATTGTTACTGAGAGAGTGTGGCTCTGCTGAA 60
Db 56 AAGGCCCATGAGATGAGTGAACCAANTTTTCTGAGAGAGTGTGGCTCTGCTGAA 115
QY 61 AGTGAACGAAAAAGTGTGCTAATGATAGCCGACATTTTGAATACATATACATCC 120
Db 116 AGTGAACGAAAAAGTGTGCTAATGATAGCCGACATTTTGAATACATATACATCC 175
QY 121 CACATTTTGAAGGCATTAATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 180
Db 176 CACATTTTGAAGGCATTAATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 235
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCAGCATTCAGAGAAACATAAGTTGATCAT 240
```


QY 578 ATACGTGCGAAGCCGAGCTTTATCCCGAGTCTCATTTCTGCGTGCCTGTGAAATG 637
DB 975 ATACGTGCGAAGCCGAGCTTTATCCCGAGTCTCATTTCTGCGTGCCTGTGAAATG 1034
QY 638 ACAGCTTTTGTGAAAAATTTTGCCTGTTGAGCAAAATCAGTAGATTTTATTGAGAAAG 697
DB 1035 ACAGCTTTTGTGAAAAATTTTGCCTGTTGAGCAAAATCAGTAGATTTTATTGAGAAAG 1094
QY 698 CAAAAGCTTCCAAATGATGATGTTTCTAGTGCATGTTTAACTGAGATCTCCGCTCCGCA 757
DB 1095 CAAAAGCTTCCAAATGATGATGTTTCTAGTGCATGTTTAACTGAGATCTCCGCTCCGCA 1154
QY 758 CCATGCTATGCTCATCATATGAAGAGATGAGCAATGCTTTAGTAGAATGAGTTAAGAT 817
DB 1155 CCATGCTATGCTCATCATATGAAGAGATGAGCAATGCTTTAGTAGAATGAGTTAAGAT 1214
QY 818 TTGTGAAAGAAAAAGCCTACTATATCTTCAAACTTCAATTTTCTGGCCAACTCTG 877
DB 1215 TTGTGAAAGAAAAAGCCTACTATATCTTCAAACTTCAATTTTCTGGCCAACTCTG 1274
QY 878 ACTATGAGAAAGATTAGAACCAAGCTGAGCATGAGGCGCAAGCAAACTCAAGC 937
DB 1275 ACTATGAGAAAGATTAGAACCAAGCTGAGCATGAGGCGCAAGCAAACTCAAGC 1334
QY 938 TTGTCACCTGAGAGAGCCAAATGAACCTGCTGCTGCTGCTGAGAGGCTGAGCAAGAAA 997
DB 1335 TTGTCACCTGAGAGAGCCAAATGAACCTGCTGCTGCTGCTGAGAGGCTGAGCAAGAAA 1394
QY 998 GCGAGAGCGCCCTCAGTTCACCTGCTGCGCACTCTGCTAATCTTCAAGAGAGAGCAAAA 1057
DB 1395 GCGAGAGCGCCCTCAGTTCACCTGCTGCGCACTCTGCTAATCTTCAAGAGAGAGCAAAA 1454
QY 1058 GGGCGCTGATCCCCGCAAGCTGCCCCAGAGCTGCCAGCGTGCAGCCGTGCTGTTAGAG 1117
DB 1455 GGGCGCTGATCCCCGCAAGCTGCCCCAGAGCTGCCAGCGTGCAGCCGTGCTGTTAGAG 1514
QY 1118 ACAGCCGCTGATGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGAGAA 1177
DB 1515 ACAGCCGCTGATGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGAGAA 1574
QY 1178 GCAATAGCTCAAGCGTCTCTCTCTGAGATATCAAAATCAGTTTCAATTTAGCCAGCA 1237
DB 1575 GCAATAGCTCAAGCGTCTCTCTCTGAGATATCAAAATCAGTTTCAATTTAGCCAGCA 1634
QY 1238 TGGCAGATCTTCAATGCTTCTCTCTCATCAGAGATGCTTTGGAATCTAATAAATCTT 1297
DB 1635 TGGCAGATCTTCAATGCTTCTCTCTCATCAGAGATGCTTTGGAATCTAATAAATCTT 1694
QY 1298 CCACTACTGATGAGGAGCAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1357
DB 1695 CCACTACTGATGAGGAGCAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1754
QY 1358 AGCAGATCTCCGAAACAGTCTGATAGAGAGAAAGCCAGATCTCCCAAGAGCTGAGCA 1417
DB 1755 AGCAGATCTCCGAAACAGTCTGATAGAGAGAAAGCCAGATCTCCCAAGAGCTGAGCA 1814
QY 1418 CCGCAGAGCTTCAAGACGAG 1477
DB 1815 CCGCAGAGCTTCAAGACGAG 1874
QY 1478 GCAGCGCCAGAGAGTCCCTTTATCTCCAGTGCATGAGAGAGAGAGAGAGAGAGAGAG 1537
DB 1875 GCAGCGCCAGAGAGTCCCTTTATCTCCAGTGCATGAGAGAGAGAGAGAGAGAGAGAG 1934
QY 1538 ACCACACAGCTTCTTTTCCGCTTTCACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAG 1597
DB 1935 ACCACACAGCTTCTTTTCCGCTTTCACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAG 1994
QY 1598 GCGTGGGCTTAAAGGCTGAGCTGAGATATCTTTGGCCCCCAGAGCTTCACTTCCCTTCC 1657
DB 1995 GCGTGGGCTTAAAGGCTGAGCTGAGATATCTTTGGCCCCCAGAGCTTCACTTCCCTTCC 2054

QY 1658 TGACCAAGCAGCTGATTTTGGCCAGAGTCTCACTTCTTACTTGTGCTCAGGCATCT 1717
DB 2055 TGACCAAGCAGCTGATTTTGGCCAGAGTCTCACTTCTTACTTGTGCTCAGGCATCT 2114
QY 1718 ACAGAGCAGTCCAGTCTTACTTGTGCTCAGCTGAGCAGCTGAGCAGCTTCCAGAGACC 1777
DB 2115 ACAGAGCAGTCCAGTCTTACTTGTGCTCAGCTGAGCAGCTGAGCAGCTTCCAGAGACC 2174
QY 1778 AAGCTATTTCTGTCGCGAGCGCGAGAAAGCAAGTGAAGAGCTGAGCTGCGCGAGCT 1837
DB 2175 AAGCTATTTCTGTCGCGAGCGCGAGAAAGCAAGTGAAGAGCTGAGCTGCGCGAGCT 2234
QY 1838 GGCATGAAGAGAGCCCTTTGAAAGCAGTTTAAAGCAGAGCTGCCAAATGGAATTG 1897
DB 2235 GGCATGAAGAGAGCCCTTTGAAAGCAGTTTAAAGCAGAGCTGCCAAATGGAATTG 2294
QY 1898 GAGAGAGCATCTGTCAGAGAACAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1957
DB 2295 GAGAGAGCATCTGTCAGAGAACAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2354
QY 1958 CTAGCTTTTCCGCGAGCATGGAATCATTTAGAGTCTCC 1995
DB 2355 CTAGCTTTTCCGCGAGCATGGAATCATTTAGAGTCTCC 2392
RESULT 14
US-10-072-012-257
Sequence 257, Application US/10072012
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Zernusen, Bryan
APPLICANT: Patnirajan, Meera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Padigaru, Murali'hara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taubier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsebrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07

Db 1972 AGGTCACGGGAGAGCTGGGGAAAGTGGCAGTCTAGCTTTTGGGCAGCATGAA 2031
QY 1981 ATCATTTAGGTCCTCC 1995
Db 2032 ATCATTTAGGTCCTCC 2046

RESULT 15

US-10-072-012-255

; Sequence 255, Application US/10072012
; Publication No. US20040033493A1

GENERAL INFORMATION:

APPLICANT: Tchiernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Patturajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 255
LENGTH: 2200
TYPE: DNA
ORGANISM: Homo sapiens
US-10-072-012-255

Query Match 59.2%; Score 1182; DB 13; Length 2200;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 814 AGATTTGTGAAGAAAAAGACCTAATAATCTCCAACTTCAATTTCTGGGCCAACTC 873
Db 919 AGATTTGTGAAGAAAAAGACCTAATAATCTCCAACTTCAATTTCTGGGCCAACTC 978
QY 874 CTGACCTATGAGAAAGATTAAAGACCAAGATCAGAGATCAGGGCCAAAGAGCAAACTC 933
Db 979 CTGACCTATGAGAAAGATTAAAGACCAAGATCAGAGATCAGGGCCAAAGAGCAAACTC 1038
QY 934 AAGCTGCTGCACCTGGAAGAACCAATGAACTGTCCTGCTGCTCAGAGGGTGGACAG 993
Db 1039 AAGCTGCTGCACCTGGAAGAACCAATGAACTGTCCTGCTGCTCAGAGGGTGGACAG 1098
QY 994 AAAAGCGAGACGCCCTCAGTCCACCTGTGCGCAGCTGTGCTACTCCAGAGCAGACAGA 1053
Db 1099 AAAAGCGAGACGCCCTCAGTCCACCTGTGCGCAGCTGTGCTACTCCAGAGCAGACAGA 1158
QY 1054 CAAAGGCCGTGTCATCCCGCAGCGTGCACAGCGTCCAGCGTGCAGCGTGCCTGTTA 1113
Db 1159 CAAAGGCCGTGTCATCCCGCAGCGTGCACAGCGTGCACAGCGTGCAGCGTGCCTGTTA 1218
QY 1114 GAGGACACCGCGTGTGTCAGAGCGTCAAGTGGGCTGACCTGTCCGACAGACAGCTGGAA 1173
Db 1219 GAGGACACCGCGTGTGTCAGAGCGTCAAGTGGGCTGACCTGTCCGACAGACAGCTGGAA 1278
QY 1174 GACAGCAATTAAGCTCAAGCTGTTCTCTCTGATATCAATCAATGATTCAATTACAGCC 1233
Db 1279 GACAGCAATTAAGCTCAAGCTGTTCTCTCTGATATCAATCAATGATTCAATTACAGCC 1338
QY 1234 AGCATGCGAGATCTTATCATGAGCTTCTCTCATCAGAAAGCTTTGGAATACATCAAA 1293
Db 1339 AGCATGCGAGATCTTATCATGAGCTTCTCTCATCAGAAAGCTTTGGAATACATCAAA 1398
QY 1294 CTTTCACTACTCTGGAAGGGAACCAAGATTAAGCCAGTTCTCCCTGTTCAGAACTA 1353
Db 1399 CTTTCACTACTCTGGAAGGGAACCAAGATTAAGCCAGTTCTCCCTGTTCAGAACTA 1458
QY 1354 TCGAGCAGACTCCCGAAACAGTCTGTATGAGAGAGAGCAGACATCCCAAGAGCTG 1413
Db 1459 TCGAGCAGACTCCCGAAACAGTCTGTATGAGAGAGAGCAGACATCCCAAGAGCTG 1518
QY 1414 CAGACCGCGCAGGCTTTCAGACAGCCAGAGCAGATGATGCTGTCAGAACAGCAGAC 1473
Db 1519 CAGACCGCGCAGGCTTTCAGACAGCCAGAGCAGATGATGCTGTCAGAACAGCAGAC 1578
QY 1474 AGTGGCACCGGCCAGAGTCCCTTTATCTCCTCACTGATGAGATGGAGAGGTGGAGAC 1533
Db 1579 AGTGGCACCGGCCAGAGTCCCTTTATCTCCTCACTGATGAGATGGAGAGGTGGAGAC 1638
QY 1534 AATTACCAACACAGCTTCTTTTGGCTTTCACAGCAGACAGCAGCTCAGAGAGTCT 1593
Db 1639 AATTACCAACACAGCTTCTTTTGGCTTTCACAGCAGACAGCAGCTCAGAGAGTCT 1698
QY 1594 GCTGGCTGGGCTTTAAGGGCTGAGACTGGATATCTTGGACCCCGCAGACCTTACCCCT 1653
Db 1699 GCTGGCTGGGCTTTAAGGGCTGAGACTGGATATCTTGGACCCCGCAGACCTTACCCCT 1758
QY 1654 TCCCTGACCGACAGCTGTATTTTGGCAGAGGCTCACAATTATCTGTGCTAGGCC 1713
Db 1759 TCCCTGACCGACAGCTGTATTTTGGCAGAGGCTCACAATTATCTGTGCTAGGCC 1818
QY 1714 ATCTACGAGGAGCTGTCAGTACTGTGCTTACAGCTGACAGCAGCTGACCTTGGGGA 1773
Db 1819 ATCTACGAGGAGCTGTCAGTACTGTGCTTACAGCTGACAGCAGCTGACCTTGGGGA 1878
QY 1774 GACCAAGTCTATTTCTGTGGCAGGCGCAGAGCCAAAGTGAACAGAGTGACTTGGCGCG 1833
Db 1879 GACCAAGTCTATTTCTGTGGCAGGCGCAGAGCCAAAGTGAACAGAGTGACTTGGCGCG 1938
QY 1834 AGCTGCGATGAAGAGAGCGCCCTTTGAAGACATTTTAAAGCGAAGCTGCCAAATGGA 1893
Db 1939 AGCTGCGATGAAGAGAGCGCCCTTTGAAGACATTTTAAAGCGAAGCTGCCAAATGGA 1998
QY 1894 TTGTGAGAGAGCATCATGTGACAGAAACAGGTCAAGGGAAGAGCTGGGGAAGTGGGAGT 1953

Db	1999	TTTGGAGAGAGCATGTCAGAGAACAGGTCACGGAGAGCTGGGAAAGTGGCAGT	2058
Qy	1954	CAGCTTAGCTTTTCGGGACAGCATGAAATCATTGAGTCTCC	1995
Db	2059	CAGCTTAGCTTTTCGGGACAGCATGAAATCATTGAGTCTCC	2100

Search completed: June 22, 2004, 10:19:07
Job time : 3262 secs

This Page Blank (uspto)

— 21 — (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2004, 02:56:17 ; Search time 172 Seconds
(without alignments)
6436.785 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995

Sequence: 1 atggcccatgagatgatgtg.....tggaatcatgaggtctcc 1995

Scoring table: OLIGO_NTC

Gapop 60.0, Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 92308

Minimum DB seq length: 1419

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1893	94.9	1998	4 US-09-816-494-3	Sequence 3, App1
2	1893	94.9	3544	4 US-09-816-494-1	Sequence 1, App1
3	19	1.0	1830	4 US-09-557-921-1	Sequence 1, App1
4	19	1.0	2170	4 US-08-961-527-312	Sequence 312, App
5	19	1.0	2377	4 US-09-920-668-3	Sequence 3, App1
6	19	1.0	2387	4 US-08-705-477E-100	Sequence 100, App
7	19	1.0	2653	2 US-08-325-553-1	Sequence 1, App1
8	19	1.0	2653	2 US-08-394-152A-1	Sequence 1, App1
9	19	1.0	2653	4 US-08-705-477E-1	Sequence 1, App1
10	19	1.0	2957	2 US-08-394-152A-48	Sequence 48, App1
11	19	1.0	3017	2 US-08-394-152A-39	Sequence 39, App1
12	19	1.0	3017	4 US-08-705-477E-39	Sequence 39, App1
13	19	1.0	3017	4 US-08-705-477E-128	Sequence 128, App
14	19	1.0	3077	4 US-08-705-477E-90	Sequence 90, App1
15	18	0.9	3083	4 US-08-961-527-138	Sequence 138, App
16	18	0.9	3549	4 US-09-008-097-5	Sequence 5, App1
17	18	0.9	4046	1 US-07-793-961A-1	Sequence 1, App1
18	18	0.9	4046	1 US-08-240-357-1	Sequence 1, App1
19	18	0.9	5852	4 US-09-853-768-10	Sequence 10, App1
20	18	0.9	7037	4 US-09-853-768-3	Sequence 3, App1
21	18	0.9	7900	4 US-08-956-171E-138	Sequence 138, App
22	18	0.9	9751	3 US-09-238-303-7	Sequence 7, App1
23	18	0.9	9751	4 US-09-946-239-7	Sequence 7, App1
24	18	0.9	19250	4 US-08-961-527-35	Sequence 35, App1
25	18	0.9	29485	4 US-09-785-381-6	Sequence 6, App1
26	18	0.9	1830121	4 US-09-557-884-1	Sequence 1, App1
27	18	0.9	1830121	4 US-09-643-990A-1	Sequence 1, App1

C 28	17	0.9	1509	4 US-09-540-236-1860	Sequence 1860, App
C 29	17	0.9	1527	4 US-09-489-039A-6404	Sequence 6404, App
C 30	17	0.9	1532	4 US-09-459-133-1	Sequence 1, App1
C 31	17	0.9	1533	1 US-08-489-733-2	Sequence 2, App1
C 32	17	0.9	1533	2 US-08-993-581B-2	Sequence 2, App1
C 33	17	0.9	1533	4 US-09-134-078-8	Sequence 8, App1
C 34	17	0.9	1596	4 US-09-484-970B-148	Sequence 148, App
C 35	17	0.9	1703	4 US-09-220-132-59	Sequence 59, App1
C 36	17	0.9	1856	2 US-08-360-606B-29	Sequence 29, App1
C 37	17	0.9	2187	4 US-09-127-219B-2	Sequence 2, App1
C 38	17	0.9	2710	4 US-09-620-312D-829	Sequence 829, App
C 39	17	0.9	2789	4 US-09-907-794A-338	Sequence 338, App
C 40	17	0.9	2789	4 US-09-905-125A-338	Sequence 338, App
C 41	17	0.9	2789	4 US-09-902-775A-338	Sequence 338, App
C 42	17	0.9	3175	4 US-09-023-655-1185	Sequence 1185, App
C 43	17	0.9	3250	3 US-08-617-860B-1	Sequence 1, App1
C 44	17	0.9	3846	2 US-08-845-161A-5	Sequence 5, App1
C 45	17	0.9	3846	3 US-09-270-751-5	Sequence 5, App1

ALIGNMENTS

RESULT 1					
US-09-816-494-3					
Sequence 3, Application US/09816494					
Patent No. 6664089					
GENERAL INFORMATION:					
APPLICANT: Meyers, Rachel A.					
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY					
FILE REFERENCE: 10448-030002					
CURRENT APPLICATION NUMBER: US/09/816,494					
PRIORITY FILING DATE: 2001-03-23					
PRIOR APPLICATION NUMBER: US 60/191,858					
PRIORITY FILING DATE: 2000-03-24					
NUMBER OF SEQ ID NOS: 10					
SOFTWARE: FastSeq for Windows Version 4.0					
SEQ ID NO 3					
LENGTH: 1998					
TYPE: DNA					
ORGANISM: Homo sapiens					
US-09-816-494-3					
Query Match					
Best Local Similarity 94.9%; Score 1893; DB 4; Length 1998;					
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
QY	1				
DB	1	ATGGCCCATGAGATGATGGAATCTCAATTTACTGAGAGGTGGCTGCTGGAA	60		
QY	61	AGTGAACGGAAGAAAGTGTCTTAATTGATGAGCCGCAATTTGTGAAATACATATCC	120		
DB	61	AGTGAACGGAAGAAAGTGTCTTAATTGATGAGCCGCAATTTGTGAAATACATATCC	120		
QY	121	CACATTTGGAAGCCATTAATCACTGCTCAACCTTATGAAGGAAAGTTGCAACG	180		
DB	121	CACATTTGGAAGCCATTAATCACTGCTCAACCTTATGAAGGAAAGTTGCAACG	180		
QY	181	GACAAAGTTAATTAACAGAGCTCATCCGACATTCAGGAAACATTAAGGTGATGAT	240		
DB	181	GACAAAGTTAATTAACAGAGCTCATCCGACATTCAGGAAACATTAAGGTGATGAT	240		
QY	241	TGACATCAGAGGTTAGTTATGATCAATCAAGCTCCCAAGATTTCTCTCTTCA	300		
DB	241	TGACATCAGAGGTTAGTTATGATCAATCAAGCTCCCAAGATTTCTCTCTTCA	300		
QY	301	GACTGTTTCTCACTGATCTTCTGGGTAACCTGAGAGAGCTTCACTCTGTCACCTG	360		
DB	301	GACTGTTTCTCACTGATCTTCTGGGTAACCTGAGAGAGCTTCACTCTGTCACCTG	360		
QY	361	CTTGACAGGTGGGTTTCTGATGTTCTCTGTTTCTCCCTGCTGTGAAGAAATATC	420		

Db 361 CTTGACAGTGGGTTGGCTGAGTCTCTGTTGTTCCCTGGCCCTGTGGAAGAAATCC 420
QY 421 ACTAGTCCCTACCTGATTTCTCAGCCTTGCTTACTGTGTGCAACATTTGGCCACC 480
Db 421 ACTTAAGTCCCTACCTGATTTCTCAGCCTTGCTTACTGTGTGCAACATTTGGCCACC 480
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCAGGAGATGTCTCAACAAGAGCTGATA 540
Db 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCAGGAGATGTCTCTCAACAAGAGCTGATG 540
QY 541 CAGAGAAATGGAGTTGGTTATGTGTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTT 600
Db 541 CAGAGAAATGGAGTTGGTTATGTGTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTT 600
QY 601 ATCCCGAGTCTCAATTTCTGCGGTGCTGTGAAATGACAGCTTTTGTGAAATTTTG 660
Db 601 ATCCCGAGTCTCAATTTCTGCGGTGCTGTGAAATGACAGCTTTTGTGAAATTTTG 660
QY 661 CCGTGTGGACAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCAAATGAGTGTGT 720
Db 661 CCGTGTGTGGACAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCAAATGAGTGTGT 720
QY 721 CTATGTGACATGTTTAACTGGGATCTCCCGTCCGCCACATGCGCTATCCCTACATCATG 780
Db 721 CTATGTGACATGTTTAACTGGGATCTCCCGTCCGCCACATGCGCTATCCCTACATCATG 780
QY 781 AAGAGATGACATGTTTAACTGGGATCTCAAGATTTGTGAAAGAAAGAAAGAAAGAAAG 840
Db 781 AAGAGATGACATGTTTAACTGGGATCTCAAGATTTGTGAAAGAAAGAAAGAAAGAAAG 840
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGAAATTAAGAAC 900
Db 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGAAATTAAGAAC 900
QY 901 CAGACTGGAGCATCAGGGCCAAAGAGCAATCTAACCTGTGCACTGTGAGAAAGCAAT 960
Db 901 CAGACTGGAGCATCAGGGCCAAAGAGCAATCTAACCTGTGCACTGTGAGAAAGCAAT 960
QY 961 GAACTGTCTCTGCTCTCAGAGGGTGGACAGAAAGAGAGAGCGCCCTCAGTCCACC 1020
Db 961 GAACTGTCTCTGCTCTCAGAGGGTGGACAGAAAGAGAGAGCGCCCTCAGTCCACC 1020
QY 1021 TGTGCGCATCTCTGCTCTCAGAGGGTGGACAGAAAGAGAGCGCCCTCAGTCCACC 1080
Db 1021 TGTGCGCATCTCTGCTCTCAGAGGGTGGACAGAAAGAGAGCGCCCTCAGTCCACC 1080
QY 1081 CCGAGGTCGCCAGGCTGAGCGCTGTGAGAGAGAGAGCGCCCTGATCCCGCAGCGTG 1140
Db 1081 CCGAGGTCGCCAGGCTGAGCGCTGTGAGAGAGAGAGCGCCCTGATCCCGCAGCGTG 1140
QY 1141 AGTGGGCTGACCTGTCTGCGCAGAGAGGCTGAGAGAGAGATTAAGGCTTCTTTC 1200
Db 1141 AGTGGGCTGACCTGTCTGCGCAGAGAGGCTGAGAGAGAGATTAAGGCTTCTTTC 1200
QY 1201 TCTCTGAGATTAATCAATCAGTTTCAATTTCAAGCAGATGAGCAGATCTTTCATGAGCTTC 1260
Db 1201 TCTCTGAGATTAATCAATCAGTTTCAATTTCAAGCAGATGAGCAGATCTTTCATGAGCTTC 1260
QY 1261 TCTCTATCAGAAAGATGCTTTTGAATACTAACAACTTTCATCCTGTGAGTGGAGCCAAAC 1320
Db 1261 TCTCTATCAGAAAGATGCTTTTGAATACTAACAACTTTCATCCTGTGAGTGGAGCCAAAC 1320
QY 1321 AAGCTATGCGAGTTTCTCCCTGTTCAGGAACCTATGAGAGCAGATCTCCCGAAACAGTCTCT 1380
Db 1321 AAGCTATGCGAGTTTCTCCCTGTTCAGGAACCTATGAGAGCAGATCTCCCGAAACAGTCTCT 1380
QY 1381 GATTAAGAGAGAGCCAGCATCCCGAAAGAGCTGAGAGCGCCAGCTTTCAGACAGCCAG 1440
Db 1381 GATTAAGAGAGAGCCAGCATCCCGAAAGAGCTGAGAGCGCCAGCTTTCAGACAGCCAG 1440
QY 1441 AGCAAGCATGTCATTCGCTGAGAAACGAGCAGAGTGCACCGCCAGAGAGTCTCCCTTTTA 1500
Db 1441 AGCAAGCATGTCATTCGCTGAGAAACGAGCAGAGTGCACCGCCAGAGAGTCTCCCTTTTA 1500

Db 1441 AGCAAGCATGTCATTCGCTGAGAAACGAGCAGAGTGCACCGCCAGAGAGTCTCCCTTTTA 1500
QY 1501 TCTTCACTGATCGAAGTGGAGCGGTGAGAGCAATTAACACACAGCTTCTTTTTCGCGC 1560
Db 1501 TCTTCACTGATCGAAGTGGAGCGGTGAGAGCAATTAACACACAGCTTCTTTTTCGCGC 1560
QY 1561 CTTTCCACAGCCACAGACACTCAAGAGTGTGTGCTGTGGGCTTTAAGGGCTGGCAC 1620
Db 1561 CTTTCCACAGCCACAGACACTCAAGAGTGTGTGCTGTGGGCTTTAAGGGCTGGCAC 1620
QY 1621 TCGGATATCTGGGCCCCCAGACCTCTACCCCTTCTGACAGCAGAGCTGTGATTTTTC 1680
Db 1621 TCGGATATCTGGGCCCCCAGACCTCTACCCCTTCTGACAGCAGAGCTGTGATTTTTC 1680
QY 1681 ACAGAGTCTCTACATTTCTTCTGTGCTGACGATCTTACGAGAGCAGTGCAGTTACTCT 1740
Db 1681 ACAGAGTCTCTACATTTCTTCTGTGCTGACGATCTTACGAGAGCAGTGCAGTTACTCT 1740
QY 1741 GCTTACAGCTGACAGCTGCGCCACTTGGCGGAGACCAAGTCTATTCTGTGCGAGCGCG 1800
Db 1741 GCTTACAGCTGACAGCTGCGCCACTTGGCGGAGACCAAGTCTATTCTGTGCGAGCGCG 1800
QY 1801 CAGAAAGCAAGTGAAGAGCTGACTGCGCGGAGCTGGCATGAGAGAGCCCTTTGAA 1860
Db 1801 CAGAAAGCAAGTGAAGAGCTGACTGCGCGGAGCTGGCATGAGAGAGCCCTTTGAA 1860
QY 1861 AAGCATTTAAACCGCAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTGAGAGAC 1920
Db 1861 AAGCATTTAAACCGCAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTGAGAGAC 1920
QY 1921 AGGTCAAGGAGAGCTGGGAAAGTGGGAGTCACTGAGTCTTGTGGGAGCATGAGAA 1980
Db 1921 AGGTCAAGGAGAGCTGGGAAAGTGGGAGTCACTGAGTCTTGTGGGAGCATGAGAA 1980
QY 1981 ATCAATGAGTCTCC 1995
Db 1981 ATCAATGAGTCTCC 1995

RESULT 2

US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589) ... (2583)
US-09-816-494-1

Query Match 94.9%; Score 1893; DB 4; Length 3544;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCATGAGATGATGGAATGGAATTTGTTACTGAGAGGTTGGTGGCTGCTGGAA 60
Db 589 ATGGCCATGAGATGATGGAATGGAATTTGTTACTGAGAGGTTGGTGGCTGCTGGAA 648
QY 61 AGTGAACGGAAGAGTGTGCTAATTTGATAGCGGCAATTTGTGGAATACATATCC 120
Db 61 AGTGAACGGAAGAGTGTGCTAATTTGATAGCGGCAATTTGTGGAATACATATCC 120

Db 649 AGTGGACGAAAAAGTCTGCTAATTGATAGCCGGCCTATTGTGAAATACATACATCC 708
Qy 121 GCAATTTGGAGCCATTATATCACTGCTCCAGCTTATGAGGGAAGTTGGACAG 180
Db 709 CACATTTTGGAGCCATTATATCACTGCTCCAGCTTATGAGGGAAGTTGGACAG 768
Qy 181 GACAAAGTTTATTCAGAGCTCATCCAGCATTCAGCAAAATTAAGTTGACATTGAT 240
Db 769 GACAAAGTTTATTCAGAGCTCATCCAGCATTCAGCAAAATTAAGTTGACATTGAT 828
Qy 241 TGCAGTCAGAGGTTGATGTTTACATCAAGAGCTCCCAAGATGTTGCTCTCTTCA 300
Db 829 TGCAGTCAGAGGTTGATGTTTACATCAAGAGCTCCCAAGATGTTGCTCTCTTCA 888
Qy 301 GACTGTTTCTCACTGATCTTCTGGGTAATCGGAGAAAGCTTCACTCTGTCACCTG 360
Db 889 GACTGTTTCTCACTGATCTTCTGGGTAATCGGAGAAAGCTTCACTCTGTCACCTG 948
Qy 361 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTGTTCCCTGGCCTCTGTGAAGAAATCC 420
Db 949 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTGTTCCCTGGCCTCTGTGAAGAAATCC 1008
Qy 421 ACTGATGCTCACTGATCTTCTGAGCTTCTGCTTCACTGTCGCAAAATGGGCCAAC 480
Db 1009 ACTGATGCTCACTGATCTTCTGAGCTTCTGCTTCACTGTCGCAAAATGGGCCAAC 1068
Qy 481 CGAATTTCTCCCAATTTTATCTTGGCTGCGCAGAGATGTCCTCAACAGAGCTGATA 540
Db 1069 CGAATTTCTCCCAATTTTATCTTGGCTGCGCAGAGATGTCCTCAACAGAGCTGATA 1128
Qy 541 CAGCAAGATGGGATGTTATGTTAAATGACAGCTATACGTCGCAAAAGCTGACTT 600
Db 1129 CAGCAAGATGGGATGTTATGTTAAATGACAGCTATACGTCGCAAAAGCTGACTT 1188
Qy 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 660
Db 1189 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 1248
Qy 661 CCGTGTGTTGACAAATCAGTATGTTTCAATTTGAGAAAGCAAAAGCTTCAATGATGTT 720
Db 1249 CCGTGTGTTGACAAATCAGTATGTTTCAATTTGAGAAAGCAAAAGCTTCAATGATGTT 1308
Qy 721 CTAGTGCATGTTTACGAGATCTCCGCTCCGCAATGCTATGCTATCATCATG 780
Db 1309 CTAGTGCATGTTTACGAGATCTCCGCTCCGCAATGCTATGCTATCATCATG 1368
Qy 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 840
Db 1369 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGTGAAGAAAGAAAGCTTACT 1428
Qy 841 ATATCTCCAACTTCAATTTTCTGGGCCAATCTGAGCTATGAGAAAGATTAAAGAC 900
Db 1429 ATATCTCCAACTTCAATTTTCTGGGCCAATCTGAGCTATGAGAAAGATTAAAGAC 1488
Qy 901 CAGACTGAGACATCAAGGCGCAAAAGCAAACTCAACTGCTGACCTGGAAGGCCAAAT 960
Db 1489 CAGACTGAGACATCAAGGCGCAAAAGCAAACTCAACTGCTGGAAGGCCAAAT 1548
Qy 961 GAACCTGTCTCTGCTCTCAGAGGTTGACAGAAAGAGAGAGCCCTCAGTCCACCC 1020
Db 1549 GAACCTGTCTCTGCTCTCAGAGGTTGACAGAAAGAGAGAGCCCTCAGTCCACCC 1608
Qy 1021 TGTGCGGACTCTGCTAATCTCAGAGGAGAGCAAAAGCCGCTGATCCCGCAGCGTG 1080
Db 1609 TGTGCGGACTCTGCTAATCTCAGAGGAGAGCAAAAGCCGCTGATCCCGCAGCGTG 1668
Qy 1081 CCCAGGTTGCGGAGGAGAGGCGGCTGTTGAGAGACAGCCGCTGATGAGAGGCGTC 1140
Db 1669 CCCAGGTTGCGGAGGAGAGGCGGCTGTTGAGAGACAGCCGCTGATGAGAGGCGTC 1728
Qy 1141 AGTGGCTGACCTGTCTCCGACAGAGGCTGAGAGACAGCAATTAAGCTCAAGCTTCTTC 1200
Db 1729 AGTGGCTGACCTGTCTCCGACAGAGGCTGAGAGACAGCAATTAAGCTCAAGCTTCTTC 1788

Qy 1201 TCTTGATATCAATCAAGTTTATATTGAGGAGATGAGAGATCTTATGAGCTTC 1260
Db 1789 TCTTGATATCAATCAAGTTTATATTGAGGAGATGAGAGATCTTATGAGCTTC 1848
Qy 1261 TCTTGATATCAATCAAGTTTATATTGAGGAGATGAGAGATCTTATGAGCTTC 1320
Db 1849 TCTTGATATCAATCAAGTTTATATTGAGGAGATGAGAGATCTTATGAGCTTC 1908
Qy 1321 AAGCTATGCAAGTTTCTCCCTGTTTCAAGAACTTATGAGAGACATCTCCGAAACCACTCT 1380
Db 1909 AAGCTATGCAAGTTTCTCCCTGTTTCAAGAACTTATGAGAGACATCTCCGAAACCACTCT 1968
Qy 1381 GATTAAGGAGAGAGAGAGATCCCAAGAGCTGAGACCGGAGGCTTCAAGAGCCAG 1440
Db 1969 GATTAAGGAGAGAGAGAGATCCCAAGAGCTGAGACCGGAGGCTTCAAGAGCCAG 2028
Qy 1441 AGCAAGCATGATGCTTCTGTCAGAAACAGAGAGATGAGAGAGCTTCTTCTTTA 1500
Db 2029 AGCAAGCATGATGCTTCTGTCAGAAACAGAGAGATGAGAGAGCTTCTTCTTTA 2088
Qy 1501 TCTTCACTGATGAGAGTGGAGGCTGAGAGCAATTTACACACAGCTTCTTTTGGC 1560
Db 2089 TCTTCACTGATGAGAGTGGAGGCTGAGAGCAATTTACACACAGCTTCTTTTGGC 2148
Qy 1561 CTTCACAGAGAGAGAGAGCTTCAAGAGTGTGCTGCTGAGGCTTAAAGGCTGGCAC 1620
Db 2149 CTTCACAGAGAGAGAGAGCTTCAAGAGTGTGCTGCTGAGGCTTAAAGGCTGGCAC 2208
Qy 1621 TCGATATCTTGGGCTCCAGAGCTTCACTGCTGAGAGAGAGAGCTTCTTCTTGGC 1680
Db 2209 TCGATATCTTGGGCTCCAGAGCTTCACTGCTGAGAGAGAGAGCTTCTTCTTGGC 2268
Qy 1681 ACAGAGTCTCACTTCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 2269 ACAGAGTCTCACTTCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2328
Qy 1741 GCTTACAGTGGAG 1800
Db 2329 GCTTACAGTGGAG 2388
Qy 1801 CAG 1860
Db 2389 CAG 2448
Qy 1861 AAGAGTTTAAAG 1920
Db 2449 AAGAGTTTAAAG 2508
Qy 1921 AGTCAAGGAG 1980
Db 2509 AGTCAAGGAG 2568
Qy 1981 ATCATTGAGGCTCC 1995
Db 2569 ATCATTGAGGCTCC 2583

RESULT 3
US-09-557-921-1
; Sequence 1, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Lucite, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.416
; CURRENT APPLICATION NUMBER: US/09/557,921
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1830

TYPE: DNA
ORGANISM: Homo sapiens
US-09-557-921-1

Query Match
Best Local Similarity 1.0%; Score 19; DB 4; Length 1830;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 TCCCGCTCCGCCACCATCG 763
Db 1323 TCCCGCTCCGCCACCATCG 1341

RESULT 4
US-08-961-527-312
Sequence 312, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ. ID NO: 312:
SEQUENCE CHARACTERISTICS:
LENGTH: 2170 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-312

Query Match
Best Local Similarity 1.0%; Score 19; DB 4; Length 2170;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 456 ACCTGTGCCACATTTGGG 474
Db 580 ACCTGTGCCACATTTGGG 598

RESULT 5
US-09-920-668-3
Sequence 3, Application US/09920668
Patent No. 6482644
GENERAL INFORMATION:
APPLICANT: Lex M. Cowsett
APPLICANT: Brett P. Monia
TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 8 EXPRESSION
FILE REFERENCE: RTS-0246

CURRENT APPLICATION NUMBER: US/09/920,668
CURRENT FILING DATE: 2001-08-01
NUMBER OF SEQ ID NOS: 49
SEQ ID NO 3
LENGTH: 2377
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (135) ... (2012)
US-09-920-668-3

Query Match
Best Local Similarity 1.0%; Score 19; DB 4; Length 2377;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 ATCGCCTACATCATGAGA 784
Db 906 ATCGCCTACATCATGAGA 924

RESULT 6
US-08-705-477E-100
Sequence 100, Application US/08705477E
Patent No. 6569432
GENERAL INFORMATION:
APPLICANT: Israeli, Ron S.
APPLICANT: Heston, Warren D.W.
APPLICANT: Fair, William R.
APPLICANT: Overfelli, Ouathak
APPLICANT: Pinto, John
TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND USES THEREOF
FILE REFERENCE: 1769/41426-G
CURRENT APPLICATION NUMBER: US/08/705,477E
CURRENT FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn version 3.1
SEQ ID NO 100
LENGTH: 2387
TYPE: DNA
ORGANISM: Homo sapiens
US-08-705-477E-100

Query Match
Best Local Similarity 1.0%; Score 19; DB 4; Length 2387;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 AGATGTGCTCTCTCTCT 297
Db 35 AGATGTGCTCTCTCTCT 53

RESULT 7
US-08-325-553-1
Sequence 1, Application US/08325553
Patent No. 5538866
GENERAL INFORMATION:
APPLICANT: Israeli, Ron S.
APPLICANT: Heston, Warren D.W.
APPLICANT: Fair, William R.
TITLE OF INVENTION: THE PROSTATE-SPECIFIC MEMBRANE ANTIGEN
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,553
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/973,337A
FILING DATE: 05 NOV 1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 1747/41426
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2653 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Carcinoma
IMMEDIATE SOURCE:
CLONE: Prostate-Specific Membrane Antigen
FEATURE:
NAME/KEY: CDS
LOCATION: 262..2511
US-08-325-553-1

Query Match 1.0%; Score 19; DB 1; Length 2653;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 AGATGTTGCCTCTCTCTCT 297
Db 35 AGATGTTGCCTCTCTCTCT 53

RESULT 8
US-08-394-152A-1
Sequence 1, Application US/08394152A
Patent No. 5935818
GENERAL INFORMATION:
APPLICANT: Israel, Ron S.
APPLICANT: Heaton, Warren D.W.
APPLICANT: Fair, William R.
TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM 330 466 DX2
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,152A
FILING DATE: 24-FEB-95
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.

REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41426-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2653 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Carcinoma
IMMEDIATE SOURCE:
CLONE: Prostate-Specific Membrane Antigen
FEATURE:
NAME/KEY: CDS
LOCATION: 262..2511
US-08-394-152A-1

Query Match 1.0%; Score 19; DB 2; Length 2653;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 AGATGTTGCCTCTCTCTCT 297
Db 35 AGATGTTGCCTCTCTCTCT 53

RESULT 9
US-08-705-477E-1
Sequence 1, Application US/08705477E
Patent No. 6569432
GENERAL INFORMATION:
APPLICANT: Israel, Ron S.
APPLICANT: Heaton, Warren D.W.
APPLICANT: Fair, William R.
APPLICANT: Overfell, Quatek
APPLICANT: Pinto, John
TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND USES THEREOF
FILE REFERENCE: 1769/41426-G
CURRENT APPLICATION NUMBER: US/08/705,477E
CURRENT FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 2653
TYPE: DNA
ORGANISM: Homo sapiens
US-08-705-477E-1

Query Match 1.0%; Score 19; DB 4; Length 2653;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 AGATGTTGCCTCTCTCTCT 297
Db 35 AGATGTTGCCTCTCTCTCT 53

RESULT 10
US-08-394-152A-48/C
Sequence 48, Application US/08394152A
Patent No. 5935818
GENERAL INFORMATION:
APPLICANT: Israel, Ron S.
APPLICANT: Heaton, Warren D.W.
APPLICANT: Fair, William R.
TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND

```
/ TITLE OF INVENTION: USES THEREOF
/ NUMBER OF SEQUENCES: 48
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Cooper & Dunham LLP
/ STREET: 1185 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: United States of America
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM 330 466 DX2
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/394,152A
/ FILING DATE: 24-FEB-95
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P.
/ REGISTRATION NUMBER: 28,678
/ REFERENCE/DOCKET NUMBER: 41426-B
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 278-0400
/ TELEFAX: (212) 391-0525
/ INFORMATION FOR SEQ ID NO: 48:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2957 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHEICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Homo Sapien
/ TISSUE TYPE: Carcinoma
/ IMMEDIATE SOURCE:
/ CLONE: Prostate Specific Membrane Antigen
/ US-08-394-152A-48

Query Match          1.0%; Score 19; DB 2; Length 2957;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 AGATGTGCTCTCTCTCT 297
DB 499 AGATGTGCTCTCTCTCT 481

RESULT 11
US-08-394-152A-39
/ Sequence 39, Application US/08394152A
/ Patent No. 5935818
/ GENERAL INFORMATION:
/ APPLICANT: Israeli, Ron S.
/ APPLICANT: Heston, Warren D.W.
/ APPLICANT: Fair, William R.
/ TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND
/ TITLE OF INVENTION: USES THEREOF
/ NUMBER OF SEQUENCES: 48
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Cooper & Dunham LLP
/ STREET: 1185 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: United States of America
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM 330 466 DX2
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
```

```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/394,152A
/ FILING DATE: 24-FEB-95
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P.
/ REGISTRATION NUMBER: 28,678
/ REFERENCE/DOCKET NUMBER: 41426-B
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 278-0400
/ TELEFAX: (212) 391-0525
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3017 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHEICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Homo Sapien
/ TISSUE TYPE: Carcinoma
/ IMMEDIATE SOURCE:
/ CLONE: Prostate Specific Membrane Antigen
/ US-08-394-152A-39

Query Match          1.0%; Score 19; DB 2; Length 3017;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 AGATGTGCTCTCTCTCT 297
DB 2519 AGATGTGCTCTCTCTCT 2537

RESULT 12
US-08-705-477E-39
/ Sequence 39, Application US/08705477E
/ Patent No. 6569432
/ GENERAL INFORMATION:
/ APPLICANT: Israeli, Ron S.
/ APPLICANT: Heston, Warren D.W.
/ APPLICANT: Fair, William R.
/ APPLICANT: Overfelli, Quathek
/ TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND USES THEREOF
/ FILE REFERENCE: 1769/41426-G
/ CURRENT APPLICATION NUMBER: US/08/705,477E
/ CURRENT FILING DATE: 1996-08-29
/ NUMBER OF SEQ ID NOS: 128
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 39
/ LENGTH: 3017
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-08-705-477E-39

Query Match          1.0%; Score 19; DB 4; Length 3017;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 AGATGTGCTCTCTCTCT 297
DB 2519 AGATGTGCTCTCTCTCT 2537

RESULT 13
US-08-705-477E-128
/ Sequence 128, Application US/08705477E
/ Patent No. 6569432
/ GENERAL INFORMATION:
/ APPLICANT: Israeli, Ron S
```

APPLICANT: Heston, Warren D.W.
APPLICANT: Fair, William R.
APPLICANT: Overfelli, Ouathek
APPLICANT: Pinto, John
TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND USES THEREOF
FILE REFERENCE: 1769/41426-G
CURRENT APPLICATION NUMBER: US/08/705,477E
CURRENT FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 128
LENGTH: 3017
TYPE: DNA
ORGANISM: Homo sapiens
US-08-705-477E-128

Query Match 1.0%; Score 19; DB 4; Length 3017;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Indels 0; Gaps 0;

QY 279 AGATGTTGCTCTCTCTCT 297
DB 2519 AGATGTTGCTCTCTCTCT 2537

RESULT 14
US-08-705-477E-90
Sequence 90, Application US/08/05477E
Patent No. 6569432
GENERAL INFORMATION:
APPLICANT: Israel, Ron S
APPLICANT: Heston, Warren D.W.
APPLICANT: Fair, William R.
APPLICANT: Overfelli, Ouathek
APPLICANT: Pinto, John
TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND USES THEREOF
FILE REFERENCE: 1769/41426-G
CURRENT APPLICATION NUMBER: US/08/705,477E
CURRENT FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 90
LENGTH: 3077
TYPE: DNA
ORGANISM: Homo sapiens
US-08-705-477E-90

Query Match 1.0%; Score 19; DB 4; Length 3077;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 AGATGTTGCTCTCTCTCT 297
DB 2579 AGATGTTGCTCTCTCTCT 2597

RESULT 15
US-08-961-527-138/C
Sequence 138, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunach
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 138:
SEQUENCE CHARACTERISTICS:
LENGTH: 3083 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-138

Query Match 0.9%; Score 18; DB 4; Length 3083;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 951 GAAGCCAATGAACCTGT 968
DB 748 GAAGCCAATGAACCTGT 731

Search completed: June 22, 2004, 09:24:34
Job time: 178 secs

This Page Blank (cspfo)

Db 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGlnTyAsnHisSer 40
QY 121 CACATTTTGAAGCCATTAAATCAACCTGCTCAAGCTTATGAAAGGAGTTGCAAG 180
Db 41 HisIleLeuGlnAlaIleAsnIleAsnCySerIleuMetLysArgArgLeuGlnGln 60
QY 181 GACAAAGTGTAAATACAGAGCTCATCCAGATTGAGGAAACATTAAGTTGACATTGAT 240
Db 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
QY 241 TGCAGTCAGAAAGTTGATTAGATTACGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 300
Db 81 CysSerGlnLysValValValValLysArgLysSerGlnAspValAlaSerLeuSerSer 100
QY 301 GACGTGTTTCCACGCTCTCTGAGTAACTGGTAACATGGAAGAGCTTCACTGTTACCTG 360
Db 101 AspCysPheLeuThrValLeuLeuGlnLysLeuGlnLysSerPheAsnSerValHisLeu 120
QY 361 CTTCAGAGTGGGTTTGCTGAGTCTCTGTTGTTTCCCTGGCTCTGTAAGAAATATCC 420
Db 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuCysGlnGlyLysSer 140
QY 421 ACTTAAGTCCCTACCTGATTTCTCAAGCTTGTACCTGTTGCAACATTTGGGCCAACC 480
Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
QY 481 CGAATTCCTCCCAATCTTATCTTGGCTCCAGGAGAGATGCTTCAACAGAGCTGATA 540
Db 161 ArgIleLeuProAsnLeuIleuGlyCysGlnArgAspValLeuAsnLysGlnLeuIle 180
QY 541 CAGCAGAAATGGATGGTATTATGTTAATGCCAGTATACCTGTCCAAAGCTGACTTT 600
Db 181 GlnGlnAsnGlyIleGlyIleValLeuAsnAlaSerThrCysProLysProAspPhe 200
QY 601 ATCCCGAGTCTCATTTCCCTGCGCTGCTGCTGTAATGACAGCTTTTGTGAAAATTTTG 660
Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
QY 661 CCGTGTGTGGAACAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCCAATGGAATGTT 720
Db 221 ProThrLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCysVal 240
QY 721 CTAGTGAAGCTGTTAGCTGGGATCTCCCGCTCCGCCACATGCGTATCGCTACATCATG 780
Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleMet 260
QY 781 AAGAGATGAGACATGCTTAGATGAAGCTTACAGATTGTGAAAGAAAAGAAAGACTTACT 840
Db 261 LysArgMetAspMetSerLeuAspGlnAlaTyArgPheValLysGlnLysArgProThr 280
QY 841 ATATCTCAAACTTCAATTTTCTGGGCCAACTCTGACTATGAGAAAGATTAAAGAAC 900
Db 281 IleSerProAsnPheAsnPheLeuGlnGlnLeuLeuAspTyrlLysLysIleLeuAsn 300
QY 901 CAGACTGGAGCATCAGAGCGCAAGAGCAAACTAACCTGCTGACCTGAGAGCCCAAT 960
Db 301 GlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisIleuGlnLysProAsn 320
QY 961 GAACCTGTCTGCTGCTCTCAGAGGTTGAGACAGAAAAGGAGAGCGCCCTCAGTCCACC 1020
Db 321 GluProValProAlaValSerGlnGlyGlnLysSerGlnThrProLeuSerProPro 340
QY 1021 TGTGCGGACTCTGCTACCTCAGAGGAGAGAGCAAAAGGCCGCTGATCCCGCAGCGTG 1080
Db 341 CysAlaAspSerAlaThrSerGlnAlaAlaGlnArgProValHisProAlaSerVal 360
QY 1081 CCCAGGCTCCAGAGCGCTGAGCGCTGCTGTTAGAGAGACAGCCCGCTGATACAGCGCTC 1140
Db 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
QY 1141 AGTGAGCTGCACCTGTCTCGCAGACAGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200

Db 381 SerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerPhe 400
QY 1201 TCTTGGAATCAAAATCAGTTTTCATTTACAGCAGAGATGGCAGCATCCCTTACATGGCTTC 1260
Db 401 SerLeuAspIleLysSerValSerTySerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
QY 1261 TCTTCATCAGAAAGATCTTTGGAAATCTAACAACTTCCACTACTCTGATGGAGCAAC 1320
Db 421 SerSerSerGlnAspAlaLeuGlnTyrlTyrlLysProSerThrThrLeuAspGlyHisAsn 440
QY 1321 AAGCTATGGCAAGTTCCTCCCTGTTACAGAACTATGAGAGCAACATCTCCGAAACAGTCTCT 1380
Db 441 LysLeuCysGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
QY 1381 GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGACAGACCGCAGGCTTACAGACGCCAG 1440
Db 461 AspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGln 480
QY 1441 AGCAGCCGATTTGCTTCGCTCAGAAACAGACAGCAGTGGACCGCCAGAGTCCCTTTTA 1500
Db 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
QY 1501 TCTCCAGTCGATCGAAGTGGGAGCGTGGAGAGCAATTAACACACAGCTTCTTTCCGC 1560
Db 501 SerProLeuHisArgSerGlySerValGlnAspAsnTyrlHisThrSerPheLeuPheGly 520
QY 1561 CTTTCACACGACGACAGCACCTCAAGAAGTCTGCTGAGCTTGAAGAGCTGAGCAGC 1620
Db 521 LeuSerThrSerGlnGlnHisIleuThrLysSerAlaGlyLeuGlnLysGlyTyrlHis 540
QY 1621 TCGGATATCTGGGCCCCCAACACCTTACCCCTTCCCTGACACAGCAGCTGTATTTGGC 1680
Db 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrlPheAla 560
QY 1681 ACAGAGTCCCTCAGACTTCTGCTGCTCAGCCATCTACAGAGAGAGAGGAGTTAGCTT 1740
Db 561 ThrGlnSerSerHisPheTySerAlaSerAlaIleTyrlGlySerAlaSerTyrlSer 580
QY 1741 GCCTACAGCTGACAGCAGCTGCCCATTTGCGGAGACCAAGTCTATTTCTGCGCAGCGG 1800
Db 581 AlaTyrlSerCysSerGlnLeuProThrCysGlyAspGlnValTyrlSerValArgArgArg 600
QY 1801 CAGAAAGCAAGTGAACAGAGCTGACCTCGGCGGAGGCTGGATGAAGAGAGAGCCCTTTGAA 1860
Db 601 GlnLysProSerAspArgAlaAspSerArgArgSerTyrlHisGlnGlnSerProPheGln 620
QY 1861 AAGCATTTAAACGACAGAGCTGCCAATGGAATTTGGAGAGAGCATCATGTGAGAGAAC 1920
Db 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGGTACAGGAGAGAGCTGGGGGAAAGTGGGAGTCAGTCTAGCTTTTGGGACAGATGAA 1980
Db 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCATTGAGGTCTCC 1995
Db 661 IleIleGlnValSer 665

RESULT 2
ABRS2407
ID ABRS2407 standard; protein; 665 AA.
XX ABRS2407;
DT 19-JUN-2003 (first entry)
XX
XX Protein relating to the invention SEQ ID NO: 148.
DE
XX anti-proliferative; hepatotropic; nephrotropic; antiarrhythmic;
XX anti-ischaemic; cardiatic; cyostatic; gene therapy; liver disease;
KW proliferative disorder; renal failure; cardiovascular disorder;
KW immunological disorder; arthritis; psoriasis; congenital heart defect;
KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.

XX Homo sapiens.
 OS
 XX
 PN WO200257460-A2.
 XX
 PD 25-JUL-2002.
 XX
 PP 20-DEC-2001; 2001MO-US050459.
 XX
 PR 20-DEC-2001; 2000US-0256668P.
 PR 30-MAR-2001; 2001US-0280186P.
 PR 01-MAY-2001; 2001US-0287735P.
 PR 05-JUN-2001; 2001US-0295648P.
 PR 25-JUN-2001; 2001US-0300465P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PI Jackson DG, Feder J, Nelson T, Muntzer G, Ramanathan C, Lee L;
 PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
 PI Kyurek S, Mcatee P, Suchard S, Banas D;
 DR WPI; 2002-559721/64.
 XX N-PSDB; ACC60572.
 XX
 PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 XX
 PS Disclosure; Fig 19; 801pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid comprising a
 CC polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotides fully defined in the specification. The polynucleotide of
 CC the invention has antiproliferative, hepatocytic, nephrotropic,
 CC antitumor, antiproliferative, cardiac, and cytostatic activity. The
 CC polynucleotide may have a use in gene therapy. A polynucleotide or
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention
 XX
 SQ Sequence 665 AA:
 Alignment Scores:
 Pred. No.: 0 Length: 665
 Score: 665.00 Matches: 665
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 5 Gaps: 0
 US-10-029-345a-108_COPY_538_2532 (1-1995) x ABR52407 (1-665)
 QY 1 ATGGCCCATGATGATGGAACCTCAATTTGTAAGAGAGTTGCTGCTGCGAA 60
 Db 1 MetAlaHisGluMetIleGlyThrGlnIleValThrIuArgLeuValAlaLeuGlu 20
 QY 61 AGTGAACGGAAGAGTCTGCTAATGATAGCCGCCATTGTTGGTAATCAATACATCC 120
 Db 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyraThrSer 40
 QY 121 CACATTTGGAAGCATTAATATCACTGCTCCAGCTTATGAGGAGGAGGTCAGAG 180
 Db 41 HisIleuGluValAlaIleAsnIleAsnCysSerIuValuValuValuValuValu 60
 QY 181 GACAAAGTGTAAATACAGAGCTCATCCAGATTCAAGCAAAACATTAAGTTGACATTGAT 240
 Db 61 AspIuValLeuIleThrGluLeuIleGlnHisSerIuValuValuValuValuValu 80

QY 241 TGCAGTACAGAGTTGATAGTATACGATCAAGACCTCCAGATGTTGCTCTCTTCA 300
 Db 81 CysSerGlnIuValValValuValuValuValuValuValuValuValuValuValu 100
 QY 301 GACTGTTTCTCAGTGTACTTCTGGGTAACCTGAGAGAGGCTTCACTCTGTTACCTG 360
 Db 101 AspCysPheLeuThrValLeuLeuGluGlyuValuValuValuValuValuValu 120
 QY 361 CTTCGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
 Db 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGluGlyuValu 140
 QY 421 ACTCTAGTCCCTACCGCATTTCTCAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 QY 481 CGAATTCCTCCAAATCTTTATCTTGGCTGCCAGAGATGCTCTCAAGAGAGCTGATA 540
 Db 161 ArgIleLeuProAsnLeuThrLeuGluGlyCysGlnArgAspValLeuLeuValuLeu 180
 QY 541 CAGCAGAAATGGAGTTGTTATGTTAAATGCCAGTATACCTGTGCCAAAGCTGACTTT 600
 Db 181 GlnGlnAsnGlyIleGlyTyraValLeuAsnAlaSerTyraThrCysProIuProAspPhe 200
 QY 601 ATCCCGAGTTCATTTCTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Db 201 IleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIuIleLeu 220
 QY 661 CCGTGGTTGAGCAAAATCAGTATGATTTCAATTGAGAAAGCAAAAGCTTCAATGGATGTT 720
 Db 221 ProThrLeuAspIuSerValAspPheIleGluValuValuValuValuValuValu 240
 QY 721 CTATGCTCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyriIleMet 260
 QY 781 AAGAGATGAGACATGCTTATGATGAGGCTTACAGATTTGTGAAAGAAAGAAAGCTTACT 840
 Db 261 LysArgMetAspPheSerLeuAspGlnAlaTyraArgPheValuValuValuValuValu 280
 QY 841 ATATCTCCAAATCTTCAATTTTCTGAGCAACTCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 281 IleSerProAsnPheAsnIleLeuGluGlnIleLeuAspTyraIuValuValuValuValu 300
 QY 901 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Db 301 GlnThrGlyAlaSerGlyProIuSerIuValuValuValuValuValuValuValu 320
 QY 961 GAACCTGCTCCCTGCTCTCAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 Db 321 GluProValProAlaValSerGluGlyGlnIuValuValuValuValuValuValuValu 340
 QY 1021 TGTGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Db 341 CysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSerVal 360
 QY 1081 CCCAGCTGCCAGCTGAG 1140
 Db 361 ProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaLeu 380
 QY 1141 AGTGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 Db 381 SerGlyLeuHisIleuSerAlaAspArgLeuGlnAspSerAsnIuValuValuValuValu 400
 QY 1201 TCTCTGATATCAATCAGTTTCATATTCAGCCAGCATGAGCATCTTACATGCTTCC 1260
 Db 401 SerLeuAspIleValSerValSerTyraSerAlaSerMetAlaAlaSerLeuHisGlyPhe 420
 QY 1261 TCTCATCATGAGAGAGCTTTGGAATCTAACAACTTCCATCTCTGATGGAGACCAAC 1320
 Db 421 SerSerSerGluuAspAlaLeuGluTyraTyraProSerThrThrLeuAspGlyThrAsn 440
 QY 1321 AAGCTATGCCAGTTCCTCCCTGTTCAAGAACTATCGAGAGAGCTCCGAAACCAAGCTCT 1380

Db 441 |||||LysLeuCySGlnPheSerProValGlnGlnLeuSerGlnGlnThrProGlnThrSerPro 460
QY 1381 GATTAGGAGGAGGAGGAGGATCCCCAAGAAAGCTGCAGACCCGAGGCTTTCAGACAGCCAG 1440
Db 461 |||||AspLysGlnGlnValAsnSerIleProLysLeuGlnThrAlaArgProSerAspSerGln 480
QY 1441 AGCAAGCGATTGCTATTCGGTCAAGAACAGCAGAGTGGCACCGCCAGAGTCCCTTTTA 1500
Db 481 SerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu 500
QY 1501 TCTCCAGCTGCATCGAAGTGGAGCGGTGAGAGACAATTACCAACACAGCTTCTTTTCGCG 1560
Db 501 SerProLeuHisArgSerGlySerValGlnAspAsnTyrHisThrSerPheLeuPheGly 520
QY 1561 CTTTCCACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
Db 521 LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheHis 540
QY 1621 TCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGCAGCTGGTATTTTGGCC 1680
Db 541 SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAla 560
QY 1681 ACAGAGTCTTCACACTTCTTACTTCTGCTCAGCCATCTTACGAGGAGCAGTGCAGTTACTCT 1740
Db 561 ThrGlnSerSerHisAspPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSer 580
QY 1741 GCCTACAGCTGCAGCAGCAGCTGCCACCTGCTACCCCTTCCCTGACACAGCAGCTGGTATTTTGGCC 1800
Db 581 AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg 600
QY 1801 CAGAGCCAGATGACAGAGCTGACTCGCGGCGAGAGCTGCATGAGAGAGAGCCCTTTGAA 1860
Db 601 GlnLysProSerAspArgAlaAspSerArgArgSerThrHisGlnGlnSerProPheGln 620
QY 1861 AAGCAGTTTAAAGCAGAAAGCTGCCAATGGAATTTGAGAGAGCATCATGTCAAGAGAAC 1920
Db 621 LysGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlnSerIleMetSerGlnAsn 640
QY 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCAAGTCTAGCTTTTTCGGGACAGATGAGAA 1980
Db 641 ArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetGln 660
QY 1981 ATCATTGAGGTCTCC 1995
Db 661 IleIleGlnValSer 665

Search completed: June 21, 2004, 14:24:44
Job time : 100.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 14:20:20 ; Search time 34.5 Seconds

(without alignments)
11124.759 Million cell updates/sec

Title: US-10-029-345a-108_COPY_538_2532

Perfect score: 665
Sequence: 1 atggcccatgagatgatgtg.....ttggaatcatgtgagctctc 1995

Scoring table:

OLIGO	Xgapop	Xgapext	Ygapop	Ygapext	Fgapop	Fgapext	Delop	Delext
	60.0	60.0	60.0	60.0	6.0	7.0	6.0	7.0

Searched: 283366 seqs, 96191526 residues

Word size: 473

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10029345/runat_21062004_140535_16051/app_query.fasta_1.2183
-DB=PIR_78 -QMT=faetan -SUFFIX=0112.rpr -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=473 -ALIGN=50 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029345_QCGN_1_1_52@runat_21062004_140535_16051 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

No matches found

Search completed: June 21, 2004, 14:30:35
Job time : 34.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 14:19:10 ; Search time 22 Seconds

(without alignments)
9443.633 Million cell updates/sec

Title: US-10-029-345a-108_COPY_538_2532

Perfect score: 665
Sequence: 1 atggcccatgagatgatctgg.....tggaaatcatctgagctctcc 1995

Scoring table:

OLIGO	Xgapop	Xgapext
60.0	60.0	60.0
Ygapop	60.0	Ygapext
60.0	60.0	60.0
Fgapop	6.0	Fgapext
7.0	6.0	7.0
Delop	6.0	Delext
7.0		

Searched: 141681 seqs, 52070155 residues

Word size: 473

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2.1/USPTO_spool_p/US10029345/runat_21062004_140534_16025/app_query.fasta_1.2183
-DB=SwissProt 42 -QFMT=fastan -SUFFIX=0112.rsp -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=75 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=473 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029345_@CGN_1_1_22_@runat_21062004_140534_16025 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	ID	Description
No.						

No matches found						

Search completed: June 21, 2004, 14:25:35
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

No matches found

Search completed: June 21, 2004, 14:29:13
Job time : 102.5 secs

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 14:19:45 ; Search time 102.5 Seconds

(without alignments)

12282.123 Million cell updates/sec

Title: US-10-029-345a-108_COPY_538_2532

Perfect score: 665

Sequence: 1 atggcccatgagatgatg.....tggaatcatgagatctcc 1995

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1017041 seqs, 315518202 residues

Word size: 473

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10029345/runat.21062004.140534.16037/app.query.fasta_1.2183
-DB=SPREMBL.25 -OFMT=fastan -SUPR=ol12.rspc -MIMATCH=0.1 -LOOPCIT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=75 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=473 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029345_@CGN_1.1.112_@runat.21062004.140534.16037 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

SPREMBL_25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp Vertebrate: *
14: sp Unclassified: *
15: sp_rv1rus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
--------	-------	-------	-------	--------	----	----	-------------

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 21, 2004, 14:24:51 ; Search time 93.5 Seconds

(without alignments)
12047.395 Million cell updates/sec

Title: US-10-029-345a-108_COPY_538_2532

Perfect score: 665
Sequence: 1 atggccatcagatgattgg.....ttggaatcattgagctctcc 1995

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1163542 seqs, 282313646 residues

Word size: 473

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=x1p
-Q=/cgn2_1/USPRO.spool_p/US10029345/runatc_21062004_140536_16108/app_query.fasta.1.2183
-DB=Published Applications AA -QFWT=fastan -SUFFIX=0112.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=75 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=473
-ALIGN=50 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10029345_@CGN_1_1_13_@runatc_21062004_140536_16108
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications_AA:*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13:	/cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

No matches found

Search completed: June 21, 2004, 14:34:57
Job time : 93.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: June 21, 2004, 14:21:20 ; Search time 24.5 Seconds

(without alignments)
8407.658 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 665
Sequence: 1 atggcccatgagatgattgg.....tggaaatcatgaggtctcc 1995

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Word size: 473

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODEL=frame+nzp.model -DEV=xlp
-O=/cgn2_1/USPTO.spool_p/US10029345/runat_21062004_140535_16081/app_query.fasta_1.2183
-DB=Issued_Patents_AA -OPMT=fastran -SUFFIX=0112.ra1 -MINMATCH=0.1 -LOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=75 -POCALIGN=200 -THR_SCORE=quality -THR_MIN=473 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10029345_@CGN_1_1_27_@runat_21062004_140535_16081 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
No matches found				

Search completed: June 21, 2004, 14:31:36
Job time : 24.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2004, 07:13:40 ; Search time 12404 Seconds

(without alignments)
6971.086 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccatgagatgtgtg9.....tggaaatcattgaggtctcc 1995

Scoring table: OLIGO-NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 1419

Total number of hits satisfying chosen parameters: 19

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_pl: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: gb_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rtd: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_by: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1995	100.0	5450	6	AX482439	AX482439 Sequence
2	1995	100.0	5450	6	AX482439	AX482439 Sequence
3	1893	94.9	1998	6	AX260342	AX260342 Sequence
4	1893	94.9	3059	6	AX278461	AX278461 Sequence
5	1893	94.9	3496	6	AX441210	AX441210 Sequence
6	1893	94.9	3521	6	AB052156	AB052156 Homo sapi
7	1893	94.9	3544	6	AX260340	AX260340 Sequence
8	1893	94.9	3566	9	AF506796	AF506796 Homo sapi
9	1893	94.9	3766	6	AX374994	AX374994 Sequence
10	1893	94.9	4790	6	BD171157	BD171157 Novel gen
11	1893	94.9	4790	6	BD183422	BD183422 Novel gen
12	1893	94.9	4790	6	AB051487	AB051487 Homo sapi
13	1842	92.3	2732	6	AX180875	AX180875 Sequence
14	1842	92.3	3104	6	AX405700	AX405700 Sequence
15	1791	89.8	2102	6	AX713989	AX713989 Sequence
16	1791	89.8	2102	9	AK055973	AK055973 Homo sapi
17	1791	89.8	5111	6	AX482372	AX482372 Sequence
18	1696	85.0	2807	9	BC031643	BC031643 Homo sapi
19	1593	79.8	2118	6	AX099933	AX099933 Sequence

ALIGNMENTS

RESULT 1
AX482439
LOCUS AX482439
DEFINITION Sequence 108 from Patent WO02057460.
ACCESSION AX482439
VERSION AX482439.1 GI:22316984
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Todderud, C.G., Bol, D., Finger, J., Lee, L., Nelson, T., Schieven, G.,
Suchard, S., Banae, D., Bassolino, D., Feder, J., Krystek, S.,
Mcatee, P., Mintier, G., Siemers, N., Jackson, D.G. and Ramanathan, C.
TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 108 25-JUL-2002;
SQUIBB BRISTOL MYERS CO (US)
FEATURES
source
1. 5450
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
538..2535
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD44451.1"
/db_xref="GI:22316985"
/translation="MAHEMIGIOIVTERVALLLESTGEKVLIDSPFVEVNTSHILE
ALININCSKLMKRLQODKVLITELIHSKAKYDIDCSQKVVYDOSDQVSLSDC
PLTVLLGKLEKFSVHLAAGFAEFSRCPGLCEKSTLVPCISQPLPVANISPT
RILPVLNLCGCRDVLNKLILQONGIGVYNAASYCPKPFIPESHLRVPVDSFCBK
ILPWLDSVDFLEKAKANGCVLHCLAGISASATIAIYIMKRMDSIDEAYRKYKE
KRPTTSPNFPILGOLLDYERKIKNOTGASGPKSKLHLERKNEBPVAVSBCQKSE
TSLSPCADSATSEAGORPVHAPSVSPVQPSLLEDSPLVQALSGHLASDRLED
SNLKRKSFSLDIKSVISASMASLHGFSSSBDALFYRPTTLDTNKLCQSPVQK
LEQTPETSPDKKEASIPKLTQPARPSDSQKRLHSVRTSSGTARSLSLPHRGS
VEDNHTSFLPLGLSTSQHLTKSAGLGGHMSDILAPQTSPTSLTSSVYFATESHF
YSASAIYGGASAYSACGLPTCCQOVYSVRRORPSDRADSRMRBESPEKOPK
RASCQMEFGESINSESRBELKGVSGSSFSGSMELIEVS"

ORIGIN

Query Match 100.0%; Score 1995; DB 6; Length 5450;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGAACTCAAAATTGTTACTGAGAGTGTGGCTCTGCTGAA 60
Db 538 ATGGCCCATGAGATGATGAACTCAAAATTGTTACTGAGAGTGTGGCTCTGCTGAA 597
QY 61 AGTGAACGGAAGAAAGTCTGCTAAATGATAGCCGGCCATTGTTGGAAATACAAATACC 120
Db 598 AGTGAACGGAAGAAAGTCTGCTAAATGATAGCCGGCCATTGTTGGAAATACAAATACC 657
QY 121 CACATTTTGAAGCCATTATATCAACTGCTCCAAGCTTATGAGCCGAGGTTGCAACAG 180
Db 658 CACATTTTGAAGCCATTATATCAACTGCTCCAAGCTTATGAGCCGAGGTTGCAACAG 717
QY 181 GACAAAGTGTAAATTAACAGAGCTCAATCCAGATTACGCAAAACATAAGTTGACATTGAT 240
Db 718 GACAAAGTGTAAATTAACAGAGCTCAATCCAGATTACGCAAAACATAAGTTGACATTGAT 777
QY 241 TGCAGTCAGAAAGTTGATAGTTAAGATCAAAAGTCCCAAGATGTGGCTCTCTCTCTCA 300
Db 778 TGCAGTCAGAAAGTTGATAGTTAAGATCAAAAGTCCCAAGATGTGGCTCTCTCTCTCA 837
QY 301 GACTGTTTCTCACTGTAATCTGGGTAACTGAGAAAGACTTCAACTCTGTTCACTG 360
Db 838 GACTGTTTCTCACTGTAATCTGGGTAACTGAGAAAGACTTCAACTCTGTTCACTG 897
QY 361 CTGGCAGTGGGTTGGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAGGAAATACC 420
Db 898 CTGGCAGTGGGTTGGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAGGAAATACC 957
QY 421 ACTCTAGTCCCTCACTGATCTTCTCAGAGCTTGTGCTTACCTGTTGCAAACTTGGCCCAACC 480
Db 958 ACTCTAGTCCCTCACTGATCTTCTCAGAGCTTGTGCTTACCTGTTGCAAACTTGGCCCAACC 1017
QY 481 CGAATTTCTCCCAATCTTATATCTGGCTGCGCAGAGATGTCTTCAACAGAGACTGATA 540
Db 1018 CGAATTTCTCCCAATCTTATATCTGGCTGCGCAGAGATGTCTTCAACAGAGACTGATA 1077
QY 541 CACAGAAATGGAGTTGTTATGTTAAATGCCATATACCTGTCCAAAGCTGACTTT 600
Db 1078 CACAGAAATGGAGTTGTTATGTTAAATGCCATATACCTGTCCAAAGCTGACTTT 1137
QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGATGACAGCTTTTGTGAGAAATTTTG 660
Db 1138 ATCCCGAGTCTCATTTCTGCGTGTGCTGATGACAGCTTTTGTGAGAAATTTTG 1197
QY 661 CCGTGTGTGACAAATCAGTATGATTGATGAGAAAGCAAAAGCTCCAAATGATGTGT 720
Db 1198 CCGTGTGTGACAAATCAGTATGATTGATGAGAAAGCAAAAGCTCCAAATGATGTGT 1257
QY 721 CTAGTCACTGTTTACGTGGAGATCCCGGCTCGGCACATGCTATGCGCTTACATCATG 780
Db 1258 CTAGTCACTGTTTACGTGGAGATCTCCGCTCGGCACATGCTATGCGCTTACATCATG 1317
QY 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTTGAAGAAAAAGACTTCT 840
Db 1318 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTTGAAGAAAAAGACTTCT 1377
QY 841 AATATCCCAACTTCAATTTTCTGGGCCAACTCTCTGACATATGAGAAGATTAAGAAC 900
Db 1378 AATATCCCAACTTCAATTTTCTGGGCCAACTCTCTGACATATGAGAAGATTAAGAAC 1437
QY 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTATGAGAAGATTAAGAAC 960
Db 1438 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTATGAGAAGATTAAGAAC 1497
QY 961 GAACCTGTCCCTGCTGTCTCAGAGGTTGACAGAAAGAGAGAGCCCTCAGTCCACC 1020
Db 1498 GAACCTGTCCCTGCTGTCTCAGAGGTTGACAGAAAGAGAGAGCCCTCAGTCCACC 1557

QY 1021 TGTGCCGACTGTGCTACTCTCAAGAGCAGCAAGCAAAAGCCGTGCATCCGCCAGCGTG 1080
Db 1558 TGTGCCGACTGTGCTACTCTCAAGAGCAGCAAGCAAAAGCCGTGCATCCGCCAGCGTG 1617
QY 1081 CCCAGGCTCCAGAGGTGACGCGTGTGTTAGAGGACAGCCCGCTGTATAGGCGCTTC 1140
Db 1618 CCCAGGCTCCAGAGGTGACGCGTGTGTTAGAGGACAGCCCGCTGTATAGGCGCTTC 1677
QY 1141 AGTGGGCTGACCTGTCCGACAGAGGCTGAGAGACAGCAATTAAGCTCAAGGTTCTTC 1200
Db 1678 AGTGGGCTGACCTGTCCGACAGAGGCTGAGAGACAGCAATTAAGCTCAAGGTTCTTC 1737
QY 1201 TCTCTGATATTAATTAAGTTTCAATTTACGACAGCATGAGCATCTTTACATGCTTC 1260
Db 1738 TCTCTGATATTAATTAAGTTTCAATTTACGACAGCATGAGCATCTTTACATGCTTC 1797
QY 1261 TCTCATCAGAAAGATGCTTTTGAATACTACAAACCTTGCATGCTGGATGGGACCAAC 1320
Db 1798 TCTCATCAGAAAGATGCTTTTGAATACTACAAACCTTGCATGCTGGATGGGACCAAC 1857
QY 1321 AAGCTATCCAGTTCTCCCTGTTCAGAGACTATTCGAGAGACTCCCGAAACCAAGTCTT 1380
Db 1858 AAGCTATCCAGTTCTCCCTGTTCAGAGACTATTCGAGAGACTCCCGAAACCAAGTCTT 1917
QY 1381 GATTAAGAGAAAGCCAGATCCCAAGAGCTGCAAGCCGCAAGCTTTCAGACAGCCAG 1440
Db 1918 GATTAAGAGAAAGCCAGATCCCAAGAGCTGCAAGCCGCAAGCTTTCAGACAGCCAG 1977
QY 1441 AGCAAGCATTTGCAATTCGGTTCAGAAACAGCAGAGTGGCAACCGCCAGAGGTCCTTTTA 1500
Db 1978 AGCAAGCATTTGCAATTCGGTTCAGAAACAGCAGAGTGGCAACCGCCAGAGGTCCTTTTA 2037
QY 1501 TCTCCACTGCAATCGAAGTGGAGCGTGGAGAGCAATTAACAACACAGCTTCTTTTGGC 1560
Db 2038 TCTCCACTGCAATCGAAGTGGAGCGTGGAGAGCAATTAACAACACAGCTTCTTTTGGC 2097
QY 1561 CTTTCCACAGCCAGCAGCACTTCAGAGATCTGCGCTGGGCTTTAAGGCTGGCAC 1620
Db 2098 CTTTCCACAGCCAGCAGCACTTCAGAGATCTGCGCTGGGCTTTAAGGCTGGCAC 2157
QY 1621 TCGGATATCTTGGCCCCCAGACCTCTTAACCCCTTCCCTGACAGAGCTGGTATTTTGGC 1680
Db 2158 TCGGATATCTTGGCCCCCAGACCTCTTAACCCCTTCCCTGACAGAGCTGGTATTTTGGC 2217
QY 1681 ACAGATCTTCACTTCTACTGCTGACCATCTTACGAGAGCAGTGCAGTFACTCT 1740
Db 2218 ACAGATCTTCACTTCTACTGCTGACCATCTTACGAGAGCAGTGCAGTFACTCT 2277
QY 1741 GCTTACAGCTGACAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGGGCGCAGGCCG 1800
Db 2278 GCTTACAGCTGACAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGGGCGCAGGCCG 2337
QY 1801 CAGAAAGCAAGTACAGAGCTGACTCGCGGCGAGCTGACATGAAGAGCCCTTTGAA 1860
Db 2338 CAGAAAGCAAGTACAGAGCTGACTCGCGGCGAGCTGACATGAAGAGCCCTTTGAA 2397
QY 1861 AAGCAGTTTAAAGCAGAGCTGACCAATGGAATTTGAGAGAGCATCATGTCAAGAAC 1920
Db 2398 AAGCAGTTTAAAGCAGAGCTGACCAATGGAATTTGAGAGAGCATCATGTCAAGAAC 2457
QY 1921 AGGTCAAGGAGAGCTGGGGAAGTGGGCAAGTCACTAGCTTTTGGGCGAGCATGGA 1980
Db 2458 AGGTCAAGGAGAGCTGGGGAAGTGGGCAAGTCACTAGCTTTTGGGCGAGCATGGA 2517
QY 1981 ATCATTTAGGTTCC 1995
Db 2518 ATCATTTAGGTTCC 2532

RESULT 2
AX482478
LOCUS

AX482478

5450 bp

DNA

linear

PAT 16-AUG-2002

DEFINITION Sequence 147 from Patent WO02057460.
ACCESSION AX482478
VERSION AX482478.1 GI:22316999
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Toddernud,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G.,
Suchard,S., Banas,D., Baesolino,D., Feder,J., Krystek,S.,
Moate,P., Minter,G., Siemers,N., Jackson,D.G. and Ramanathan,C.
TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 147 25-JUL-2002;
SQUIBB BRISTOL MYERS CO (US)
FEATURES
source
1. 5450
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
538. 2535
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD44453.1"
/db_xref="GI:22317000"
/db_xref="REMBL:CAD44453"
/translation="MAHEMIGTIVTBRLVALLBSGTBKYLLIDSRPFVRYNTSHILE
AININCKLMKRLQODKVLITELIOHSAKVDIDCSQVYVDDSDVDASLSDC
FLTVLKLKLEKSPNSVHLAGEFAEPRCPGEGSTLVPICIPCPVANIPT
RLPNLYGCRDVLNKLKELIIONGIGVLAASYCPEDPIPESHHLYVNDSPCEK
ILPMLKSVDFIEKAVASNGCVLHCLAGISRSATLAIAYIMRMMSIDEAYRVE
KRPITSPNRFPLQQLIDYEKTIKNTGASPRSKLILHEKNEVPVAVSEGGQSE
TPSPPCADSAISBAQRPVHPASVPSPVPSLSDSLVQALSGHLSADRLSD
SNLKXGSPSLDIKSVSYASMAASLHGFSSSEDALEYKSTLIDSTNKLQSPVQ
LSQETPSTSPDKKEASIPKLOTPARPSDSQSKRLHSVRTSSGTQARSLSPHRS
VEDNHTSFLFGSTSQOHLTKSAGLGLKMHSDILAPOTSTSLTSSWTFATESHF
YSASATYGGASAYSACSQLPTCGPOVYSVRRORRDBADSRMHESPREKQPK
RSCQMEFSGINSRRELBKVGSGSSFSMSLEIYS"
ORIGIN
Query Match 100.0%; Score 1995; DB 6; Length 5450;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCCATGATGATGATGAACTCAATTGTTACTGAGAGTTGGTGGCTGTGCGGAA 60
DB 538 ATGGCCCATGATGATGATGAACTCAATTGTTACTGAGAGTTGGTGGCTGTGCGGAA 597
QY 61 AGTGAACGGAAAGTGTCTAATGATAGCCGGCATTTGTGAATACATPACATCC 120
DB 598 AGTGAACGGAAAGTGTCTAATGATAGCCGGCATTTGTGAATACATPACATCC 657
QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 180
DB 658 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 717
QY 181 GACAAAGTGAATTAATCAGAGCTCATCAGCATTTACGCGAAACATTAAGTTGACATGAT 240
DB 718 GACAAAGTGAATTAATCAGAGCTCATCAGCATTTACGCGAAACATTAAGTTGACATGAT 777
QY 241 TGACATGAGAAAGTTAGTTACGATCAAGTCCCAAGATGTTGGCTCTCTCTTCA 300
DB 778 TGACATGAGAAAGTTAGTTACGATCAAGTCCCAAGATGTTGGCTCTCTCTTCA 837
QY 301 GACTGTTTCTCACTGACTTCTGGGTAATCTGGAGAAGAGCTTCACTGTCTTCACTG 360
DB 838 GACTGTTTCTCACTGACTTCTGGGTAATCTGGAGAAGAGCTTCACTGTCTTCACTG 897
QY 361 CTGACAGTGGGTTTGTGAGTTCTCTGTTGTTTCTGAGCTCTGTGGAAGAAATCC 420
DB 898 CTGACAGTGGGTTTGTGAGTTCTCTGTTGTTTCTGAGCTCTGTGGAAGAAATCC 957
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGTCTTACTCTGTGCAACATTTGGGCAAC 480

DB 958 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGTCTTACTCTGTGCAACATTTGGGCAAC 1017
QY 481 CGAATTTCTCCCAATTTTATCTTGGCTGCGCAGCGAGATGCTCTCAACAGAGCTGATA 540
DB 1018 CGAATTTCTCCCAATTTTATCTTGGCTGCGCAGCGAGATGCTCTCAACAGAGCTGATA 1077
QY 541 CAGCAATAGGAGTTGGTATGTTAAATGACAGCTATACCTGTCCAAAGCTGACTT 600
DB 1078 CAGCAATAGGAGTTGGTATGTTAAATGACAGCTATACCTGTCCAAAGCTGACTT 1137
QY 601 ATCCCGAGTCTCATTTCCCTGGGCTGCTGCTGATGATGAGAGCTTTTGTGAGAAATTTG 660
DB 1138 ATCCCGAGTCTCATTTCCCTGGGCTGCTGCTGATGATGAGAGCTTTTGTGAGAAATTTG 1197
QY 661 CCGTGTGAGCAAAATCAGTATGATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTT 720
DB 1198 CCGTGTGAGCAAAATCAGTATGATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTT 1257
QY 721 CTAGTGCATGTTTACCTGGAGATCTCCGGCTCCGACATGCTATGCTTACATCATG 780
DB 1258 CTAGTGCATGTTTACCTGGAGATCTCCGGCTCCGACATGCTTACATCATCATG 1317
QY 781 AAGAGATGAGACATGCTTTAGATGAGCTTACAGATTTGTGAAAGAAAAAGACCTACT 840
DB 1318 AAGAGATGAGACATGCTTTAGATGAGCTTACAGATTTGTGAAAGAAAAAGACCTACT 1377
QY 841 ATATCTCCAAATTTCAATTTTCTGGGCAACTCTGAGTATGAGAAAGATTAAGAAC 900
DB 1378 ATATCTCCAAATTTCAATTTTCTGGGCAACTCTGAGTATGAGAAAGATTAAGAAC 1437
QY 901 CAGACTGAGCATCAGAGGCGCAAAAGCAAACTCAACTCTGTGACCTTGAGAGAGCCAAAT 960
DB 1438 CAGACTGAGCATCAGAGGCGCAAAAGCAAACTCAACTCTGTGAGAGAGAGCCAAAT 1497
QY 961 GAACTGTGCTGCTGCTGCTCTCAGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 1498 GAACTGTGCTGCTGCTGCTCTCAGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1557
QY 1021 TGTGCGCACTGTCTAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1558 TGTGCGCACTGTCTAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1617
QY 1081 CCCAGGTCCTCAGAGGCTGAG 1140
DB 1618 CCCAGGTCCTCAGAGGCTGAG 1677
QY 1141 AGTGGCTGCACTGTGCGCAG 1200
DB 1678 AGTGGCTGCACTGTGCGCAG 1737
QY 1201 TCTCTGATATCAATCAGTTTCAATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1738 TCTCTGATATCAATCAGTTTCAATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1797
QY 1261 TCTCTATCAGAAAGTCTTTGGAATCTAATAAATCTTCACTATCTGTGATGGAGCAAC 1320
DB 1798 TCTCTATCAGAAAGTCTTTGGAATCTAATAAATCTTCACTATCTGTGATGGAGCAAC 1857
QY 1321 AAGCTATGCAAGTTTCTCCCTGTTTCAAGAACTATGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1858 AAGCTATGCAAGTTTCTCCCTGTTTCAAGAACTATGAGAGAGAGAGAGAGAGAGAGAGAG 1917
QY 1381 GATTAAG 1440
DB 1918 GATTAAG 1977
QY 1441 AGCAAGCATGATGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 1978 AGCAAGCATGATGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2037
QY 1501 TCTCCAGTGCATGAG 1560

Db 2038 TCTCCACTGCATCGAAGTGGAGCGGTGGAGACATTACACACCAGCTTCCTTTTGGC 2097
Qy 1561 CTTTCCACCGACGACGACCTCAACGAAGTCTGTGGCCCTTGAAGGCTGGCAC 1620
Db 2098 CTTTCCACCGACGACGACCTCAACGAAGTCTGTGGCCCTTGAAGGCTGGCAC 2157
Qy 1621 TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGACTGGTATTTGGCC 1680
Db 2158 TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGACTGGTATTTGGCC 2217
Qy 1681 ACAGAGTCTCTACACTTCTATCTGTCTCAGCCATCTTACCGAGGCACTGGCCAGTTACTCT 1740
Db 2218 ACAGAGTCTCTACACTTCTATCTGTCTCAGCCATCTTACCGAGGCACTGGCCAGTTACTCT 2277
Qy 1741 GCCTACAGCTGACACGACGCTGCCCACTTGGGGAGACCAAGTCTATTTGTGGCAGAGCGG 1800
Db 2278 GCCTACAGCTGACACGACGCTGCCCACTTGGGGAGACCAAGTCTATTTGTGGCAGAGCGG 2337
Qy 1801 CAGAGGCCAAGTGAACAGAGCTGACTCGCGGGAGAGTGGCATGMAAGAGACCCCTTTGAA 1860
Db 2338 CAGAGGCCAAGTGAACAGAGCTGACTCGCGGGAGAGTGGCATGMAAGAGACCCCTTTGAA 2397
Qy 1861 AAGAGTTTAAACGCAACAGCTGCAATATTTGAGAGAGCATCATGTGAGAGAC 1920
Db 2398 AAGAGTTTAAACGCAACAGCTGCAATATTTGAGAGAGCATCATGTGAGAGAC 2457
Qy 1921 AGGTCAAGGGAAGAGTGGGGAAGTGGGAGTCACTAGCTTTTGGGGCAGCATGGA 1980
Db 2458 AGGTCAAGGGAAGAGTGGGGAAGTGGGAGTCACTAGCTTTTGGGGCAGCATGGA 2517
Qy 1981 ATCATTGAGGTCTCC 1995
Db 2518 ATCATTGAGGTCTCC 2532

RESULT 3
AX260342
LOCUS AX260342 1998 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 3 from Patent WO0173059.
ACCESSION AX260342
VERSION AX260342.1 GI:16509305
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1
Meyers, R.A.
3862 and 2117: dual specificity phosphatase molecules and uses
therefor
Patent: WO 0173059-A 3 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1. 1998
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES
source

ORIGIN

Query Match 94.9%; Score 1893; DB 6; Length 1998;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGATGATGATGAACTCAATTTGTTACTGAGAGTTGGCTGTGCTGGAA 60
Db 1 ATGGCCCATGATGATGATGAACTCAATTTGTTACTGAGAGTTGGCTGTGCTGGAA 60
Qy 61 AGTGAACGGAAGAAAGTGTCTGCTAATTTAGTACCGGCACTTTTGAAATACATATCC 120
Db 61 AGTGAACGGAAGAAAGTGTCTGCTAATTTAGTACCGGCACTTTTGAAATACATATCC 120
Qy 121 GACATTTTGAAGCCATTAAATTCAGAGTCTCCAGCTTATAGCGAAGGTTGCAACG 180
Db 121 GACATTTTGAAGCCATTAAATTCAGAGTCTCCAGCTTATAGCGAAGGTTGCAACG 180

Db 121 GACATTTTGAAGCCATTAAATTCAGAGTCTCCAGCTTATAGCGAAGGTTGCAACG 180
Qy 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGATTGACGCAAAATTAAGTTGACATTGAT 240
Db 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGATTGACGCAAAATTAAGTTGACATTGAT 240
Qy 241 TGGAGCAGAAAGTTGTACTTTACATGCAAACTCCCAAGAGTTGGCTCTCTCTTTCA 300
Db 241 TGGAGCAGAAAGTTGTACTTTACATGCAAACTCCCAAGAGTTGGCTCTCTCTTTCA 300
Qy 301 GACTGTTTTCTCACTGTAATTTCTGGGTAAACGTGAGAAAGCTTCAACTCTGTACCTG 360
Db 301 GACTGTTTTCTCACTGTAATTTCTGGGTAAACGTGAGAAAGCTTCAACTCTGTACCTG 360
Qy 361 CTTGCAAGTGGGTTTGTCTGAGTTCTCTGTGTTTCCCTGGCTCTGTGAGAAATTC 420
Db 361 CTTGCAAGTGGGTTTGTCTGAGTTCTCTGTGTTTCCCTGGCTCTGTGAGAAATTC 420
Qy 421 ACTCTAGTCCCTACCGGATTTCTGACGCTTGCTTACCTGTTGGCAATGGGGCAAC 480
Db 421 ACTCTAGTCCCTACCGGATTTCTGACGCTTGCTTACCTGTTGGCAATGGGGCAAC 480
Qy 481 CGAATTTCTCCCAATCTTATCTTGGCTGCGACGAGATGTCCTCAACAGAGCTGATA 540
Db 481 CGAATTTCTCCCAATCTTATCTTGGCTGCGACGAGATGTCCTCAACAGAGCTGATA 540
Qy 541 CAGCAAGATGGGATTTGTTATGTTTAATGCGAGCTTAACTGTCCAAAGCTGACTTT 600
Db 541 CAGCAAGATGGGATTTGTTATGTTTAATGCGAGCTTAACTGTCCAAAGCTGACTTT 600
Qy 541 CAGCAAGATGGGATTTGTTATGTTTAATGCGAGCTTAACTGTCCAAAGCTGACTTT 600
Db 541 CAGCAAGATGGGATTTGTTATGTTTAATGCGAGCTTAACTGTCCAAAGCTGACTTT 600
Qy 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGAGATGACAGCTTTTGGAGAAATTTTG 660
Db 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGAGATGACAGCTTTTGGAGAAATTTTG 660
Qy 661 CCGTGTGGAACAAATCAGTAGATTTCAATGAGAAAGCAAAAGCTCCATGATGTGTT 720
Db 661 CCGTGTGGAACAAATCAGTAGATTTCAATGAGAAAGCAAAAGCTCCATGATGTGTT 720
Qy 721 CTAGTGCATGTTTAAAGCTGGGATCTCCCGCTCCGACATCGCTATCGCTTACATCATG 780
Db 721 CTAGTGCATGTTTAAAGCTGGGATCTCCCGCTCCGACATCGCTATCGCTTACATCATG 780
Qy 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGAGAAAGAAAAAGACTTACT 840
Db 781 AAGAGATGACATGCTTTAGATGAGCTTACAGATTTGAGAAAGAAAAAGACTTACT 840
Qy 841 ATATCTCCAAATCTTCAATTTTCTGGGCCAACTCTCTGAGCTATGAGAAAGATTAAGAC 900
Db 841 ATATCTCCAAATCTTCAATTTTCTGGGCCAACTCTCTGAGCTATGAGAAAGATTAAGAC 900
Qy 901 CAGACTGAGATCAGAGGCGCAAAAGCAAACTCAAGCTGTGCACTGGAGAAACCAAT 960
Db 901 CAGACTGAGATCAGAGGCGCAAAAGCAAACTCAAGCTGTGCACTGGAGAAACCAAT 960
Qy 961 GAACTGTCCCTGTCTCTCAGAGGTTGACAGAAAGAGAGAGCGCCCTCAAGTCCACC 1020
Db 961 GAACTGTCCCTGTCTCTCAGAGGTTGACAGAAAGAGAGAGCGCCCTCAAGTCCACC 1020
Qy 1021 TGTGCCCATCTGTCTACTCAGAGGCGAGCAAAAGGCGCGGATCCGCGAGAGGTG 1080
Db 1021 TGTGCCCATCTGTCTACTCAGAGGCGAGCAAAAGGCGCGGATCCGCGAGAGGTG 1080
Qy 1081 CCGAGCGTGGCCACGCGGAGCGCGTCTGTGTAGAGACAGCCCGCTGTGATCAGGCGCTC 1140
Db 1081 CCGAGCGTGGCCACGCGGAGCGCGTCTGTGTAGAGACAGCCCGCTGTGATCAGGCGCTC 1140
Qy 1141 AGTGGCTGCACTGTCCGAGACAGGCTGGAAGAAGCAATAGCTCAAGCTTCTTC 1200
Db 1141 AGTGGCTGCACTGTCCGAGACAGGCTGGAAGAAGCAATAGCTCAAGCTTCTTC 1200
Qy 1201 TCTCTGATATCAAAATCAGTTTCAATTTAGCCAGCATGGCAGATCTTACATGGCTTC 1260
Db 1201 TCTCTGATATCAAAATCAGTTTCAATTTAGCCAGCATGGCAGATCTTACATGGCTTC 1260

QY 1261 TCCTCATCAGAAAGATGCTTTGGAATACTACAAACCTTCCACTACTGTGATGGAGCAAC 1320
 DB 1261 TCCTCATCAGAAAGATGCTTTGGAATACTACAAACCTTCCACTACTGTGATGGAGCAAC 1320
 QY 1321 AAGCTATGCGAGTTCTCCCTGTTTCAGGAACTATCGGAGCAGACTCCGAAACAGTCTT 1380
 DB 1321 AAGCTATGCGAGTTCTCCCTGTTTCAGGAACTATCGGAGCAGACTCCGAAACAGTCTT 1380
 QY 1381 GATAGAGAGAGAACCGCAATCCCAAGAACTGCAACCGCCGAGCTTTAGACAGCCAG 1440
 DB 1381 GATAGAGAGAGAACCGCAATCCCAAGAACTGCAACCGCCGAGCTTTAGACAGCCAG 1440
 QY 1441 AGCAAGGATTTGATTCGCTCAGAACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1500
 DB 1441 AGCAAGGATTTGATTCGCTCAGAACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1500
 QY 1501 TCTCACTGCATCGAAGTGGAGGAGTGGAGAGCAATTACCAACAGCTTCTTTTCGCG 1560
 DB 1501 TCTCACTGCATCGAAGTGGAGGAGTGGAGAGCAATTACCAACAGCTTCTTTTCGCG 1560
 QY 1561 CTTTCCACCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
 DB 1561 CTTTCCACCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
 QY 1621 TCGGATATCTTGGGCCCCCAGACCTCTACCCCTTCCCTGACCAAGCAGCTGTAATTTGGC 1680
 DB 1621 TCGGATATCTTGGGCCCCCAGACCTCTACCCCTTCCCTGACCAAGCAGCTGTAATTTGGC 1680
 QY 1681 ACAGAGTCTCTACACTTCTACTCTGCTCTACAGCCTACAGAGCAGAGTGCAGTACTCT 1740
 DB 1681 ACAGAGTCTCTACACTTCTACTCTGCTCTACAGCCTACAGAGCAGAGTGCAGTACTCT 1740
 QY 1741 GCTTACAGCTGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1800
 DB 1741 GCTTACAGCTGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1800
 QY 1801 CAGAGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860
 DB 1801 CAGAGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860
 QY 1861 AAGCAGTTTAAAGCAGAGAGCTGCCAATATGAAATTTGAGAGAGCATCATGTCAAGAAAC 1920
 DB 1861 AAGCAGTTTAAAGCAGAGAGCTGCCAATATGAAATTTGAGAGAGCATCATGTCAAGAAAC 1920
 QY 1921 AGGTCAAGGAGAGAGTGGGAGAAAGTGGGAGTCAAGTCAAGTCTTTTCGAGCAGATGAA 1980
 DB 1921 AGGTCAAGGAGAGAGTGGGAGAAAGTGGGAGTCAAGTCAAGTCTTTTCGAGCAGATGAA 1980
 QY 1981 ATCATTGAGGTCCTC 1995
 DB 1981 ATCATTGAGGTCCTC 1995

RESULT 4
 AX278461 3059 bp DNA linear PART 02-NOV-2001
 LOCUS AX278461
 DEFINITION Sequence 1 from Patent WO0177340.
 ACCESSION AX278461
 VERSION AX278461.1 GI:16605915
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Duecker, K.
 TITLE Identification of a dual specificity phosphatase: dusp-10
 JOURNAL Patent: WO 0177340-A 1 18-Oct-2001;
 MERCK PATENT GmbH (DE)
 FEATURES
 1. .3059
 Location/Qualifiers
 /organism="Homo sapiens"

/mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 127..2124
 /note="unnamed protein product"

CDS

/codon_start=1
 /protein_id="CAD10545.1"
 /db_xref="GI:16605916"
 /db_xref="REMBL:CAD10545"
 /translation="MAHEMIGIOIVTERVALLESTGEKVLIDSRPFVYNTSHLE
 AININCSKLMKRLQODKRLITBELIOLHSAKHVYDIDSGKVVYVYDSSQDVASLSDC
 PLIVLGLKLEKSFNSVHLIAGPAESRCPELCEKSLVLPCLISQPLVPAVNIPT
 RILPNIYLICQDVLNKELMQNGIGVYLANSTCPKPDIPESHFLRVVDSFCEK
 ILPMLDKSVDFIEKASNGCVLHCLAGISRSATIALIYIMKMEVMSIDEXYRFYKE
 KPPTSPNPNFGLDYEKIKNONGAGSPSKLHLLEKNEPVPVAVSEGCQKSE
 TPLSPCADSATSEAAQGPVPAVSPVSPVOPSLLEDSPLVIALSGHLSADRLD
 SKLKRSPSLDKSVYSASMAASLHGBSSSDALEYKRPSTLLDSTNKLCPSPVQZ
 LSEQPEPSPDKEASIPKLOTARPSDSQRLHSVRTSSGTQRLSLSPATSSHS
 VEDNYHSTSLFGLSTSOHLTASAGLGMWSDILAPQSTPLSLSSVYFATSSHF
 YASASLYGGSASVAVSCQLPTCGQVYVRRRQGTSPDRADSRMSWHEESPKEQFK
 RSCQMEFSESIMSENRSEBELKVGQSSFFSGSMELIYVS"

ORIGIN

Query Match 94.9%; Score 1893; DB 6; Length 3059;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATTTGGAATCTCAAAATTTGTAATGAGAGTGGTGGCTCTGCTGAA 60
 DB 127 ATGGCCCATGAGATGATTTGGAATCTCAAAATTTGTAATGAGAGTGGTGGCTCTGCTGAA 186
 QY 61 AGTGAACGGAAAAAGTGTGCTAATTTGATAGCCGSCATTTGTGAAATACATATACCC 120
 DB 187 AGTGAACGGAAAAAGTGTGCTAATTTGATAGCCGSCATTTGTGAAATACATATACCC 246
 QY 121 CACATTTTGAAGCCATTAAATCAACTGCTCCAACTTATGAAAGGAAAGTTGCAACAG 180
 DB 247 CACATTTTGAAGCCATTAAATCAACTGCTCCAACTTATGAAAGGAAAGTTGCAACAG 306
 QY 181 GACAAAGTGTAAATTAACAGCTCATCCAGCATTCAGCGCAACATPAAGTTGATGAT 240
 DB 307 GACAAAGTGTAAATTAACAGCTCATCCAGCATTCAGCGCAACATPAAGTTGATGAT 366
 QY 241 TGCAGTCAGAAAGTGTATGATTAACATCAAACTCCCAAGATGTTGCTCTCTCTTCA 300
 DB 367 TGCAGTCAGAAAGTGTATGATTAACATCAAACTCCCAAGATGTTGCTCTCTCTTCA 426
 QY 301 GACTGTTTTCTCACTGTACTTCTGGGTAAACTGAGAGAGCTTCAACTGTGTTCACTG 360
 DB 427 GACTGTTTTCTCACTGTACTTCTGGGTAAACTGAGAGAGCTTCAACTGTGTTCACTG 486
 QY 361 CTTCAGAGTGGTGTGCTGAGTCTCTGTTGTTTCCCTGCGCTCTGTGAGAAATTC 420
 DB 487 CTTCAGAGTGGTGTGCTGAGTCTCTGTTGTTTCCCTGCGCTCTGTGAGAAATTC 546
 QY 421 ACTCTAGTCCCTACCGCATTTCTCAGCTTGCTTCACTGTTGCCAAATTTGGGCCAAC 480
 DB 547 ACTCTAGTCCCTACCGCATTTCTCAGCTTGCTTCACTGTTGCCAAATTTGGGCCAAC 606
 QY 481 CGAATTTCTCCCAATCTTATCTTGGCTGCCAGCAGATGTCCTCAACAAGAGCTGATA 540
 DB 607 CGAATTTCTCCCAATCTTATCTTGGCTGCCAGCAGATGTCCTCAACAAGAGCTGATA 666
 QY 541 CAGCAGATGAGATGTTATGTTTAAATGCCAGTATACCTGTGCCAAAGCTGACTTT 600
 DB 667 CAGCAGATGAGATGTTATGTTTAAATGCCAGTATACCTGTGCCAAAGCTGACTTT 726
 QY 601 ATCCCGAGTCTCATTTCTGCGTGGCTGCTGGAATGACAGCTTTTGTGAGAAATTTTG 660
 DB 727 ATCCCGAGTCTCATTTCTGCGTGGCTGCTGGAATGACAGCTTTTGTGAGAAATTTTG 786
 QY 661 CCGTGGTTGAGCAAAATCAGTATGATTTCAATGAGAAAGCAAAAGCTCCAAATGATGTGT 720
 DB 787 CCGTGGTTGAGCAAAATCAGTATGATTTCAATGAGAAAGCAAAAGCTCCAAATGATGTGT 846

QY 721 CTAAGTGCATGTTAGTGGGATCTCCCGCTCCGSCACATCGCTATCGCTACATATG 780
DB 847 CTAAGTGCATGTTAGTGGGATCTCCCGCTCCGSCACATCGCTATCGCTACATATG 906
QY 781 AAGAGATGGAATGCTCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGACTACT 840
DB 907 AAGAGATGGAATGCTCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGACTACT 966
QY 841 ATATCTCAAACTTCAATTTTCTGGGCACTCCTGGAATATGAGAAAGATTAAGAAC 900
DB 967 ATATCTCAAACTTCAATTTTCTGGGCACTCCTGGAATATGAGAAAGATTAAGAAC 1026
QY 901 CAGACTGAGCATGAGGCGCAAGAGCAATCTCAAGCTGTGCACTGAGAGCAAT 960
DB 1027 CAGACTGAGCATGAGGCGCAAGAGCAATCTCAAGCTGTGCACTGAGAGCAAT 1086
QY 961 GAACCTGTCCCTGTCTCTAGAGGCTGACAGAAAAAGCAGACGCGCTCACTCCACC 1020
DB 1087 GAACCTGTCCCTGTCTCTAGAGGCTGACAGAAAAAGCAGACGCGCTCACTCCACC 1146
QY 1021 TGTGCGCATCTGTCTCTAGAGGCTGACAGAAAAAGCAGACGCGCTCACTCCACC 1080
DB 1147 TGTGCGCATCTGTCTCTAGAGGCTGACAGAAAAAGCAGACGCGCTCACTCCACC 1206
QY 1081 CCCAGCGTCCGCAAGCGTGCAGCGCTGTGTAAGAGCAGCGCGCTGTGTAAGAGC 1140
DB 1207 CCCAGCGTCCGCAAGCGTGCAGCGCTGTGTAAGAGCAGCGCGCTGTGTAAGAGC 1266
QY 1141 AGTGGGCTGCACCTGTCTCGCAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
DB 1267 AGTGGGCTGCACCTGTCTCGCAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1326
QY 1201 TCTCTGATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1260
DB 1327 TCTCTGATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1386
QY 1261 TCTCTGATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1320
DB 1387 TCTCTGATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1446
QY 1321 AAGCTATGCGAATGTTCTCCCTGTTCAGAACTATGGAAGCAGCTCCGAAACCACTCT 1380
DB 1447 AAGCTATGCGAATGTTCTCCCTGTTCAGAACTATGGAAGCAGCTCCGAAACCACTCT 1506
QY 1381 GATTAAGAGAGAGCAGCAATCCCAAGAGCTGACAGCGCGCTCAAGAGCGAG 1440
DB 1507 GATTAAGAGAGAGCAGCAATCCCAAGAGCTGACAGCGCGCTCAAGAGCGAG 1566
QY 1441 AGCAAGCAGATTGCAATCGGTCAAGAACCAAGAGAGTGGACCGCCAGAGTCCCTTTTA 1500
DB 1567 AGCAAGCAGATTGCAATCGGTCAAGAACCAAGAGAGTGGACCGCCAGAGTCCCTTTTA 1626
QY 1501 TCTCCACTGATGGAAGTGGAGCGTGGAGACATTAACAACAAGCTTCTTTTCGAG 1560
DB 1627 TCTCCACTGATGGAAGTGGAGCGTGGAGACATTAACAACAAGCTTCTTTTCGAG 1686
QY 1561 CTTTTCACGAGCAGAGCACTCAAGAGCTGAGCTGGGCTTAAGGGCTGGAGC 1620
DB 1687 CTTTTCACGAGCAGAGCACTCAAGAGCTGAGCTGGGCTTAAGGGCTGGAGC 1746
QY 1621 TCGGATATTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGAGTGTATTTTTC 1680
DB 1747 TCGGATATTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGAGTGTATTTTTC 1806
QY 1681 AAGAGTCTCTACATTTCTACTCTGCTGAGCAATGACAGAGAGTGTATTTTTC 1740
DB 1807 AAGAGTCTCTACATTTCTACTCTGCTGAGCAATGACAGAGAGTGTATTTTTC 1866
QY 1741 GCTTACAGCTGAGCAGCTGCGCACTTCCGAGAGCAAGTATTTCTGTGGAGAGGAG 1800
DB 1867 GCTTACAGCTGAGCAGCTGCGCACTTCCGAGAGCAAGTATTTCTGTGGAGAGGAG 1926

QY 1801 CAGAGCCAGTGAAGAGCTGACTCGGCGGAGAGCTGGCATTAAGAGAGCCCTTTGAA 1860
DB 1927 CAGAGCCAGTGAAGAGCTGACTCGGCGGAGAGCTGGCATTAAGAGAGCCCTTTGAA 1986
QY 1861 AAGCAGTTTAAACGAGAGAGCTGCCAAATGGAATTTGAGAGAGCATGTCAGAGAAC 1920
DB 1987 AAGCAGTTTAAACGAGAGAGCTGCCAAATGGAATTTGAGAGAGCATGTCAGAGAAC 2046
QY 1921 AGTCAAGGGAAGAGCTGGGGAAGTGGGAGTCACTGCTAGCTTTTCCGAGCATGGA 1980
DB 2047 AGTCAAGGGAAGAGCTGGGGAAGTGGGAGTCACTGCTAGCTTTTCCGAGCATGGA 2106
QY 1981 ATCATTGAGGTCTCC 1995
DB 2107 ATCATTGAGGTCTCC 2121

RESULT 5
AX441210 3496 bp DNA linear PAT 28-JUN-2002
LOCUS AX441210

DEFINITION Sequence 1 from Patent WO0226597.

ACCESSION AX441210

VERSION AX441210.1 GI:2165766

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches 1993; Conservative

0; Mismatches

2; Indels

0; Gaps

0;

1

562

QY

DB

QY

DB

QY

Db 982 ACCTAGTCCCTACATGATTTCTCAGCCCTTGCTTACCTGTGCAACATTTGGGCCAC 1041
 Qy 481 CGAATTTCTCCCATCTTTATCTTGGCTGCCAGAGATGTCTCTCAACAGAGCTGATA 540
 Db 1042 CGAATTTCTCCCATCTTTATCTTGGCTGCCAGAGATGTCTCTCAACAGAGCTGATA 1101
 Qy 541 CAGCAGAAATGGATGGTATGTTGTTAAATGCCAGCTATACCTGTCCAAAGCCGACTT 600
 Db 1102 CAGCAGAAATGGATGGTATGTTGTTAAATGCCAGCTATACCTGTCCAAAGCCGACTT 1161
 Qy 601 ATCCCGAGTCTGATTTCTGCGTGTGCTGATGATGACGCTTTGTGAAATTTTG 660
 Db 1162 ATCCCGAGTCTGATTTCTGCGTGTGCTGATGATGACGCTTTGTGAAATTTTG 1221
 Qy 661 CCGTGTGTGACAAATAGTAGATTTTATGAGAAAGCAAAAGCCTTCATGATGTGT 720
 Db 1222 CCGTGTGTGACAAATAGTAGATTTTATGAGAAAGCAAAAGCCTTCATGATGTGT 1281
 Qy 721 CTAGTGACTGTTAGCTGGGATCTCCGCTCCGCCACATCGCTATCGCTCATGATG 780
 Db 1282 CTAGTGACTGTTAGCTGGGATCTCCGCTCCGCCACATCGCTATCGCTCATGATG 1341
 Qy 781 AAGAGATGACATGTCTTATGATGAACTTAAAGATTTGTGAAAGAAAAAGACCTACT 840
 Db 1342 AAGAGATGACATGTCTTATGATGAACTTAAAGATTTGTGAAAGAAAAAGACCTACT 1401
 Qy 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGACATATGAGAAAGATTAAGAC 900
 Db 1402 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGACATATGAGAAAGATTAAGAC 1461
 Qy 901 CAGACTGAGACATCAGGGCCAAAGAGAACTAAGTGCTGACCTGGAGAGCCAAAT 960
 Db 1462 CAGACTGAGACATCAGGGCCAAAGAGAACTAAGTGCTGACCTGGAGAGCCAAAT 1521
 Qy 961 GAACCTGTCCCTGCTGTCTCAGAGGCTGACAGAAAAAGCAGAGCCCTCAGTCCACC 1020
 Db 1522 GAACCTGTCCCTGCTGTCTCAGAGGCTGACAGAAAAAGCAGAGCCCTCAGTCCACC 1581
 Qy 1021 TGTGCGACTCTGTCTCAGAGGCTGACAGAAAAAGCAGAGCCCTCAGTCCACC 1080
 Db 1582 TGTGCGACTCTGTCTCAGAGGCTGACAGAAAAAGCAGAGCCCTCAGTCCACC 1641
 Qy 1081 CCCAGGCTGCCAGCCGTGACGCTGTGATGAGAGACCCCGCTGTGATCAGGGCTC 1140
 Db 1642 CCCAGGCTGCCAGCCGTGACGCTGTGATGAGAGACCCCGCTGTGATCAGGGCTC 1701
 Qy 1141 AGTGGCTGACCTGTCCGAGACAGGCTGAAAGACAGCAATAGCTCAAGCCTTCCCTC 1200
 Db 1702 AGTGGCTGACCTGTCCGAGACAGGCTGAAAGACAGCAATAGCTCAAGCCTTCCCTC 1761
 Qy 1201 TCTCTGATATCAATCAGTTTATATTTAGCCAGCAGATGACATCTTACATGCTTC 1260
 Db 1762 TCTCTGATATCAATCAGTTTATATTTAGCCAGCAGATGACATCTTACATGCTTC 1821
 Qy 1261 TCTCTCATCAGAAATGTTTGAATACTAACAACCTTCCACTCTGTGATGGAGCCAC 1320
 Db 1822 TCTCTCATCAGAAATGTTTGAATACTAACAACCTTCCACTCTGTGATGGAGCCAC 1881
 Qy 1321 AAGCTATGACAGTTTCTCCCTGTTACAGAACTATGAGAGAGACTCCCGAAACAGTCT 1380
 Db 1882 AAGCTATGACAGTTTCTCCCTGTTACAGAACTATGAGAGAGACTCCCGAAACAGTCT 1941
 Qy 1381 GATTAAGAGAGACCCAGCATCCCAAGAGCTGACAGCCGAGCCTTACAGACGAC 1440
 Db 1942 GATTAAGAGAGACCCAGCATCCCAAGAGCTGACAGCCGAGCCTTACAGACGAC 2001
 Qy 1441 AGCAAGGATTTGATTTGGTTCAGAAACAGACAGAGTGGCACCGCCAGAGAGTCCCTTTA 1500
 Db 2002 AGCAAGGATTTGATTTGGTTCAGAAACAGACAGAGTGGCACCGCCAGAGAGTCCCTTTA 2061
 Qy 1501 TCTCAGCTGATGAGATGGAGAGTGGAGAGCAATTAACACAGCAGCTTCTTTCCGG 1560

Db 2062 TCTCAGCTGATGAGATGGAGAGCTGAGAGACAAATTAACACAGCAGCTTCTTTCCGG 2121
 Qy 1561 CTTTCACGAGCAGAGAGACCTCAGAGAGTGTGAGCTGGGCTTAAAGGCTGGAC 1620
 Db 2122 CTTTCACGAGCAGAGAGACCTCAGAGAGTGTGAGCTGGGCTTAAAGGCTGGAC 2181
 Qy 1621 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGCC 1680
 Db 2182 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGCC 2241
 Qy 1681 ACAGAGTCTTACACTTCTTCTGCTGACAGCCATATACAGAGAGAGAGAGAGAGAGAG 1740
 Db 2242 ACAGAGTCTTACACTTCTTCTGCTGACAGCCATATACAGAGAGAGAGAGAGAGAGAG 2301
 Qy 1741 GCCTAAGCTGACAGCCAGCTGCCCCAATTTGCGAGAGACCAAGTCTATTTGTGGCAGCGG 1800
 Db 2302 GCCTAAGCTGACAGCCAGCTGCCCCAATTTGCGAGAGACCAAGTCTATTTGTGGCAGCGG 2361
 Qy 1801 CAGAGCCAAAGTACAGAGCTGACTCGCGGAGCTGGCATGAGAGAGAGAGAGAGAGAGAG 1860
 Db 2362 CAGAGCCAAAGTACAGAGCTGACTCGCGGAGCTGGCATGAGAGAGAGAGAGAGAGAGAG 2421
 Qy 1861 AAGCAGTTTAAAGCAGAACTGCCAAATGAAATTTGAGAGAGATCATGTGAGAGAC 1920
 Db 2422 AAGCAGTTTAAAGCAGAACTGCCAAATGAAATTTGAGAGAGATCATGTGAGAGAC 2481
 Qy 1921 AGGTACGGGAAAGAGCTGGGAAAGTGGGAGTCACTTACCTTTTCCGACAGATGAA 1980
 Db 2482 AGGTACGGGAAAGAGCTGGGAAAGTGGGAGTCACTTACCTTTTCCGACAGATGAA 2541
 Qy 1981 ATCATGAGGTCTCC 1995
 Db 2542 ATCATGAGGTCTCC 2556

RESULT 6
 AB052156
 LOCUS Homo sapiens MKP-7 mRNA for MAPK phosphatase-7, complete cds.
 DEFINITION
 AB052156
 VERSION
 AB052156.1 GI:13548676
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 Masuda, K., Shima, H., Watanabe, M. and Kikuchi, K.
 TITLE
 MKP-7, a novel mitogen-activated protein kinase phosphatase,
 functions as a shuttle protein
 JOURNAL
 J. Biol. Chem. 276 (42), 39002-39011 (2001)
 MEDLINE
 21486429
 PUBMED
 11489891
 REFERENCE
 2 (bases 1 to 3521)
 AUTHORS
 Masuda, K., Shima, H. and Kikuchi, K.
 TITLE
 Direct Substitution
 Submitted (05-DEC-2000) Kouhei Masuda, Institute for Genetic
 Medicine, Division of Biochemical Oncology and Immunology, Kita-ku
 Kitas-5-Jou Nishi7-tyoume, Sapporo, Hokkaido 060-0815, Japan
 (E-mail: kouhei.hokudai.ac.jp, Tel:81-11-706-5536,
 Fax:81-11-707-6839)
 FEATURES
 source
 1..3521
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="12"
 /cell_line="Jurkat"
 1..3521
 /gene="MKP-7"
 564..2561
 /gene="MKP-7"
 /codon_start=1

```
/product="MAPK phosphatase-7"  
/protein_id="BAB40814.1"  
/db_xref="GI:13548677"  
/translation="MAHEMIGIQIVTERLVALLSESTEKYLIDSEPFVEYINTSHILE  
AININCKLMKRRLQODKYLITELIQHSAGHKVDIDSKVWVYDSSQDVASLSDC  
FLVLLGKLEKSFNSVHLIAFGPAFSPGCLCEKSGTLVPTCLSPVANIPT  
RIIPNLYIGCDRVLNKEIMQONGIYVLANASNTCPKPIFESHRLSPVNDSPCEK  
ILPMLKSFDEIEKAKASNGCVLHCLAGISRSATIAIYIKRMMSLIDEAYRPFKE  
KRPRTISPNRFLGQLDYREKLIKNOTGASGPKSLKLIHEKNEPVPASVSGCQSE  
TPISPCASBATEBAAGQRPVHPASVPSPVPSPLLEBSPVLOALSGHLSDRLSD  
SNLKRFSFLDIKSVSASMAASLHGFSSSDALYYKPSLTLDITNKLCPSPVQ  
LSQPEPESFLGKESASI PKKLQTPARPSDSQKRLHSVRSSGTPARSLSPVDS  
VEDNHTSFLFGLSTSQOHLTKSAGLKGWMDILAPQSTSLSSWYATSESHF  
YSASAIYGSASYSAYSCSQLPTCPDQVSVRRROKPSRADRSRMHESPEKQFK  
RSCOMEFBSIMSENRSRBELKQVSGSSFSGSMETIIVS"
```

ORIGIN

```
Query Match          94.9%; Score 1893; DB 9; Length 3521;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ATGGCCCATGAGATGATGGAACCTCAATTGTTACTGAGAGTTGGTGGCTCTGCTGAA 60  
DB ATGGCCCATGAGATGATGGAACCTCAATTGTTACTGAGAGTTGGTGGCTCTGCTGAA 623  
QY 61 AGTGAACGGAAAAAGTGTCTAATTGATAGCCGGCATTTGTGGAATACATACATCC 120  
DB AGTGAACGGAAAAAGTGTCTAATTGATAGCCGGCATTTGTGGAATACATACATCC 683  
QY 121 CACATTTTGGAAACCAATTAATCACTGCTCCAAAGTTATGAAAGCAAGTTGCAACG 180  
DB CACATTTTGGAAACCAATTAATCACTGCTCCAAAGTTATGAAAGCAAGTTGCAACG 743  
QY 181 GACAAAGTGTATTTACAGAGCTCATTCAGCATTCAGCCAAACATTAAGTTGATGAT 240  
DB GACAAAGTGTATTTACAGAGCTCATTCAGCATTCAGCCAAACATTAAGTTGATGAT 803  
QY 241 TGCAGTCAGAAAGTTGATGATTCAGATCAAAAGTCCCAAGATGTGGCTCTCTCTTCA 300  
DB TGCAGTCAGAAAGTTGATGATTCAGATCAAAAGTCCCAAGATGTGGCTCTCTCTTCA 863  
QY 301 GACTGTTTCTCACTGACTTCTGGGTAAACCTGGAAGAAGCTTCAACTCTGTTCACTG 360  
DB GACTGTTTCTCACTGACTTCTGGGTAAACCTGGAAGAAGCTTCAACTCTGTTCACTG 923  
QY 361 CTTCGAGGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAGAAATCC 420  
DB CTTCGAGGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAGAAATCC 983  
QY 421 ACTCTAGTCCCTAAGCTGATTTCTCAGCCCTTGCTTACCTGTTGCAAACTGGGCAAC 480  
DB ACTCTAGTCCCTAAGCTGATTTCTCAGCCCTTGCTTACCTGTTGCAAACTGGGCAAC 1043  
QY 481 CGAATTCCTCCCAATCTTTATCTTGGCTGACAGAGATGTCCTCAACAGAGCTGATA 540  
DB CGAATTCCTCCCAATCTTTATCTTGGCTGACAGAGATGTCCTCAACAGAGCTGATA 1103  
QY 541 CAGCAGAAATGGGATGTTGTTATGTTAAATGCCAGCTATACCTGTCCAAAAGCTGACTT 600  
DB CAGCAGAAATGGGATGTTGTTATGTTAAATGCCAGCTATACCTGTCCAAAAGCTGACTT 1163  
QY 601 ATCCCGAGTCTCATTTCCGCGGTGCGCTGGAATGACAGCTTTGTGAGAAATTTTG 660  
DB ATCCCGAGTCTCATTTCCGCGGTGCGCTGGAATGACAGCTTTGTGAGAAATTTTG 1223  
QY 1164 ATCCCGAGTCTCATTTCCGCGGTGCGCTGGAATGACAGCTTTGTGAGAAATTTTG 1223  
DB ATCCCGAGTCTCATTTCCGCGGTGCGCTGGAATGACAGCTTTGTGAGAAATTTTG 1283  
QY 1224 CGGTGTTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCCAATGATGTGTT 1283  
DB CGGTGTTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCCAATGATGTGTT 1283  
QY 1284 CTATGTCACGTGTTAGCTGGGATCTCCCGCTCCGCAACATCGCTATCCCTACATCANG 1343  
DB CTATGTCACGTGTTAGCTGGGATCTCCCGCTCCGCAACATCGCTATCCCTACATCANG 1343
```

```
QY 781 AAGAGATGAGCATGCTCTTAGATGAGCTTACAGATTTGTGAAGAAAAAGACCTACT 840  
DB AAGAGATGAGCATGCTCTTAGATGAGCTTACAGATTTGTGAAGAAAAAGACCTACT 1403  
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTCTGACATATGAGAAAGATTAAAGC 900  
DB ATATCTCCAACTTCAATTTTCTGGGCAACTCTCTGACATATGAGAAAGATTAAAGC 1463  
QY 1404 ATATCTCCAACTTCAATTTTCTGGGCAACTCTCTGACATATGAGAAAGATTAAAGC 1463  
DB ATATCTCCAACTTCAATTTTCTGGGCAACTCTCTGACATATGAGAAAGATTAAAGC 1463  
QY 901 CAGACTGAGGATCTCAGGGGCAAGAGCAAACTCAAGCTGCTGACCTGAGAAAGCAAT 960  
DB CAGACTGAGGATCTCAGGGGCAAGAGCAAACTCAAGCTGCTGACCTGAGAAAGCAAT 1523  
QY 1464 CAGACTGAGGATCTCAGGGGCAAGAGCAAACTCAAGCTGCTGACCTGAGAAAGCAAT 1523  
DB CAGACTGAGGATCTCAGGGGCAAGAGCAAACTCAAGCTGCTGACCTGAGAAAGCAAT 1523  
QY 961 GAACCTGTCCTCCTGCTCTCAGAGGATGAGCAAGAAAGAGAACCCCTCAGTCCACC 1020  
DB GAACCTGTCCTCCTGCTCTCAGAGGATGAGCAAGAAAGAGAACCCCTCAGTCCACC 1583  
QY 1524 GAACCTGTCCTCCTGCTCTCAGAGGATGAGCAAGAAAGAGAACCCCTCAGTCCACC 1583  
DB GAACCTGTCCTCCTGCTCTCAGAGGATGAGCAAGAAAGAGAACCCCTCAGTCCACC 1583  
QY 1021 TGTGCTGACTGCTACTCTCAGAGGACAGAGCAAAAGCCCTGACCTCCGACAGCTG 1080  
DB TGTGCTGACTGCTACTCTCAGAGGACAGAGCAAAAGCCCTGACCTCCGACAGCTG 1643  
QY 1584 TGTGCTGACTGCTACTCTCAGAGGACAGAGCAAAAGCCCTGACCTCCGACAGCTG 1643  
DB TGTGCTGACTGCTACTCTCAGAGGACAGAGCAAAAGCCCTGACCTCCGACAGCTG 1643  
QY 1081 CCCAGCGGCCCGAGGTCGAGCGCTGCTGTTAGAGACAGCCGCTGTTACAGCGCTTC 1140  
DB CCCAGCGGCCCGAGGTCGAGCGCTGCTGTTAGAGACAGCCGCTGTTACAGCGCTTC 1703  
QY 1644 CCCAGCGGCCCGAGGTCGAGCGCTGCTGTTAGAGACAGCCGCTGTTACAGCGCTTC 1703  
DB CCCAGCGGCCCGAGGTCGAGCGCTGCTGTTAGAGACAGCCGCTGTTACAGCGCTTC 1703  
QY 1141 AGTGGGCTGCACTGTCGACAGACAGGCTGAAAGACAGCAATTAAGCTCAAGCGTTC 1200  
DB AGTGGGCTGCACTGTCGACAGACAGGCTGAAAGACAGCAATTAAGCTCAAGCGTTC 1763  
QY 1201 TCTCTGATATCAATCAAGTTTCATATTCAGCCAGCATGACGATCCTTACATGCTTC 1260  
DB TCTCTGATATCAATCAAGTTTCATATTCAGCCAGCATGACGATCCTTACATGCTTC 1823  
QY 1261 TCTCTGATATCAATCAAGTTTCATATTCAGCCAGCATGACGATCCTTACATGCTTC 1320  
DB TCTCTGATATCAATCAAGTTTCATATTCAGCCAGCATGACGATCCTTACATGCTTC 1883  
QY 1824 TCTCTGATATCAATCAAGTTTCATATTCAGCCAGCATGACGATCCTTACATGCTTC 1883  
DB TCTCTGATATCAATCAAGTTTCATATTCAGCCAGCATGACGATCCTTACATGCTTC 1943  
QY 1321 AAGCTATGCAAGATGCTTCCCTGTTGAGAACTATGAGAGCAGACTCCGAAACAGTCT 1380  
DB AAGCTATGCAAGATGCTTCCCTGTTGAGAACTATGAGAGCAGACTCCGAAACAGTCT 1943  
QY 1381 GATTAAGAGAGAGCAGACATCCCAAGAGCTGACAGCCGACAGGCTTACAGACAGCAG 1440  
DB GATTAAGAGAGAGCAGACATCCCAAGAGCTGACAGCCGACAGGCTTACAGACAGCAG 2003  
QY 1944 GATTAAGAGAGAGCAGACATCCCAAGAGCTGACAGCCGACAGGCTTACAGACAGCAG 2003  
DB GATTAAGAGAGAGCAGACATCCCAAGAGCTGACAGCCGACAGGCTTACAGACAGCAG 2063  
QY 1441 AGCAAGCATTTGCAATTCGTCAGAACCAAGAGCAGATGAGCAACGCGCCAGAGGCTTTTA 1500  
DB AGCAAGCATTTGCAATTCGTCAGAACCAAGAGCAGATGAGCAACGCGCCAGAGGCTTTTA 2063  
QY 2004 AGCAAGCATTTGCAATTCGTCAGAACCAAGAGCAGATGAGCAACGCGCCAGAGGCTTTTA 1500  
DB AGCAAGCATTTGCAATTCGTCAGAACCAAGAGCAGATGAGCAACGCGCCAGAGGCTTTTA 2123  
QY 1501 TCTCCAGTGCATCGAAGTGGAGAGCTGAGAGCAATTAACACACAGCTTCTTTTGGAC 1560  
DB TCTCCAGTGCATCGAAGTGGAGAGCTGAGAGCAATTAACACACAGCTTCTTTTGGAC 2123  
QY 2064 TCTCCAGTGCATCGAAGTGGAGAGCTGAGAGCAATTAACACACAGCTTCTTTTGGAC 1560  
DB TCTCCAGTGCATCGAAGTGGAGAGCTGAGAGCAATTAACACACAGCTTCTTTTGGAC 2123  
QY 1561 CTTTCACACGAGCAGACAGCACTCAAGAGTGTGCTGCTGAGCTTAAAGGGCTGGCAC 1620  
DB CTTTCACACGAGCAGACAGCACTCAAGAGTGTGCTGCTGAGCTTAAAGGGCTGGCAC 2183  
QY 2124 CTTTCACACGAGCAGACAGCACTCAAGAGTGTGCTGCTGAGCTTAAAGGGCTGGCAC 2183  
DB CTTTCACACGAGCAGACAGCACTCAAGAGTGTGCTGCTGAGCTTAAAGGGCTGGCAC 1680  
QY 1621 TGGGATTCCTGGCCCCCCCCAGACTCTTACCCCTTCCCTGACCAAGAGAGCTGATTTGCC 1680  
DB TGGGATTCCTGGCCCCCCCCAGACTCTTACCCCTTCCCTGACCAAGAGAGCTGATTTGCC 2243  
QY 2184 TGGGATTCCTGGCCCCCCCCAGACTCTTACCCCTTCCCTGACCAAGAGAGCTGATTTGCC 2243  
DB TGGGATTCCTGGCCCCCCCCAGACTCTTACCCCTTCCCTGACCAAGAGAGCTGATTTGCC 1740  
QY 1681 ACAGAGTCTCTCACTTCACTCTGCTCAGCATCTACAGAGGAGAGTGGCAGTACTCT 1740  
DB ACAGAGTCTCTCACTTCACTCTGCTCAGCATCTACAGAGGAGAGTGGCAGTACTCT 2303  
QY 2244 ACAGAGTCTCTCACTTCACTCTGCTCAGCATCTACAGAGGAGAGTGGCAGTACTCT 2303  
DB ACAGAGTCTCTCACTTCACTCTGCTCAGCATCTACAGAGGAGAGTGGCAGTACTCT 1800  
QY 1741 GCTTACAGCTGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
DB GCTTACAGCTGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2363  
QY 2304 GCTTACAGCTGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2363  
DB GCTTACAGCTGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860  
QY 1801 CAGAAAGCAGAGAGAGAGTGAATCTGCGGCGGAGAGAGTGGAGAGAGAGAGAGAGAG 1860  
DB CAGAAAGCAGAGAGAGAGTGAATCTGCGGCGGAGAGAGTGGAGAGAGAGAGAGAGAG 2423  
QY 2364 CAGAAAGCAGAGAGAGAGTGAATCTGCGGCGGAGAGAGTGGAGAGAGAGAGAGAG 2423  
DB CAGAAAGCAGAGAGAGAGTGAATCTGCGGCGGAGAGAGTGGAGAGAGAGAGAGAGAG 1920  
QY 1861 AAGCAGTTTAAACGAGAAAGCTGCCAATTGAAATTGAGAGAGCATGTCAGAGAAC 1920  
DB AAGCAGTTTAAACGAGAAAGCTGCCAATTGAAATTGAGAGAGCATGTCAGAGAAC 1920
```

Db 2424 AAGCAGTTTAAAGCAGAACTCCCAATGGAATTTGAGAGAGCATCATGATGAGAAAC 2483
QY 1921 AGGTACCGGGAAGAGCTGGGGAAGTGGGAGTCACTGATCTGTTTGGGACAGATGAA 1980
Db 2484 AGGTACCGGGAAGAGCTGGGGAAGTGGGAGTCACTGATCTGTTTGGGACAGATGAA 2543
QY 1981 ATGATTGAGGTCTCC 1995
Db 2544 ATCATTGAGGTCTCC 2558

RESULT 7
AX260340
LOCUS AX260340 3544 bp DNA linear PART 26-OCT-2001
DEFINITION Sequence 1 from Patent WO0173059.
ACCESSION AX260340
VERSION AX260340.1 GI:16509303
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Meyers, R.A.
38692 and 21117: dual specificity phosphatase molecules and uses
therefor
JOURNAL Patent: WO 0173059-A 1 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifiers
1..3544
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
589..2586
/note="unassigned protein product"
/codon_start=1
/protein_id="CAD10218.1"
/db_xref="GI:16509304"
/db_xref="REMBL:CAD10218"
/translation="MAHEMIGTQIVTERVALLESGETEKLIDSRPFVEYNTSHILE
AININCSKLMKRLQODKVLITELIOHSAHKVIDIOSQVYVDDSDVVASLSDC
FLTVLIGKLEKSPNSVHLIAGAPSPSCGKSTIPVISCIPCLPVNIEPT
RIIPNLYLGGQRPVNLKELMQNGICGVANASATCKRPPIPSHRLPVNDSPECK
ILPWLKSVDFIEKARASNGCVLHCLAGISRSATTAIAIMKMDMSIDEAYRVEKE
KRPITLBNFNLQILDYEKIKINQTAGISPKSLKILHLEKNEBVPVSEGGQSE
TPSPCADATSEAGORPVHPASVPSVQSLLEDPVLOALSGHLSADRLD
SNLKRSPSIDIKSVSASMAASLHGFSSSEDALEVKRSTTLIDGNTLCPSPVOE
LSQGTPEPSPDKREASIPKLTQTPARPSDSKRLHVSRTSSSTABRSLSLPHRGS
VEDNHTSFLFGLSTSQOHLTKSAGJLKGWHSIDILAPQTSSTLSLSMTFATESHF
YSASAIYGGASYSAYSCSQLPTCGDQVSVRROROKRPSDADRSRWHEBSPKQFK
RRSCQMEFBSIMSENRSRELKGVGSQSSFSGSMELIEVS"

ORIGIN
Query Match 94.9%; Score 1893; DB 6; Length 3544;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAAGTCAATGTTACTGAGAGGTTGGTGGCTCTGCTGGAA 60
Db 589 ATGGCCCATGAGATGATGGAAGTCAATGTTACTGAGAGGTTGGTGGCTCTGCTGGAA 648
QY 61 AGTGAACGGAAGAAAGTGTCTAATTGATAGCGGCGCATTTGTGGAATACATATCCTC 120
Db 649 AGTGAACGGAAGAAAGTGTCTAATTGATAGCGGCGCATTTGTGGAATACATATCCTC 708
QY 121 CACATTTTGAAGCATTAATATCAATGCTCCAGATTATGAAGCGCAAGTTGCAACAG 180
Db 709 CACATTTTGAAGCATTAATATCAATGCTCCAGATTATGAAGCGCAAGTTGCAACAG 768
QY 181 GACAAAGTGTAAATTAAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATTGAT 240
Db 769 GACAAAGTGTAAATTAAGAGCTCATCAGCATTCAGCGAAACATAAGTTGACATTGAT 828

QY 241 TGCAGTCAGAAAGTTGTATGTTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
Db 829 TGCAGTCAGAAAGTTGTATGTTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 888
QY 301 GACTGTTTCTCAGTGTACTTCTGGGTAACTGAGAGAGCTTCAACTCTGTTCACTG 360
Db 889 GACTGTTTCTCAGTGTACTTCTGGGTAACTGAGAGAGCTTCAACTCTGTTCACTG 948
QY 361 CTTCAGAGTGGGTTTGTGATGTTCTCTGTTGTTTCCCTGGGCTCTGTGAGAGAAATCC 420
Db 949 CTTCAGAGTGGGTTTGTGATGTTCTCTGTTGTTTCCCTGGGCTCTGTGAGAGAAATCC 1008
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCGCTTGATCTGTTGCGCAAACTGGGCGCAAC 480
Db 1009 ACTCTAGTCCCTACCTGCAATTTCTCAGCGCTTGATCTGTTGCGCAAACTGGGCGCAAC 1068
QY 481 CGAATCTTCCCAATCTTATCTGAGCTGCCAGCGAGATGTCTTCAACAAGAGCTGATA 540
Db 1069 CGAATCTTCCCAATCTTATCTGAGCTGCCAGCGAGATGTCTTCAACAAGAGCTGATA 1128
QY 541 CAGCAGAAATGGGATTTGTTATGTTAAATGCCAGTATACCTGTCCAAAGCTGACTT 600
Db 1129 CAGCAGAAATGGGATTTGTTATGTTAAATGCCAGTATACCTGTCCAAAGCTGACTT 1188
QY 601 ATCCCGAGATCTCATTTCTGCGGTGCTGCTGAGATGAGACAGCTTTGTGAGAAATTTTG 660
Db 1189 ATCCCGAGATCTCATTTCTGCGGTGCTGCTGAGATGAGACAGCTTTGTGAGAAATTTTG 1248
QY 661 CCGTGTGTTGACCAATCAGTATGTTTATTGAGAAAGCAAAAGCTTCAATGATGTTGTT 720
Db 1249 CCGTGTGTTGACCAATCAGTATGTTTATTGAGAAAGCAAAAGCTTCAATGATGTTGTT 1308
QY 721 CTAGTCACTGTTTATGCTGGGATCTTCCGCTCCGCGCCAGCATGCTATGCTTACATCATG 780
Db 1309 CTAGTCACTGTTTATGCTGGGATCTTCCGCTCCGCGCCAGCATGCTATGCTTACATCATG 1368
QY 781 AAGAGATGAGCATGCTTATGATGAGCTTACAGTTTGTGAGAAAGAAAGAAAGCTTACT 840
Db 1369 AAGAGATGAGCATGCTTATGATGAGCTTACAGTTTGTGAGAAAGAAAGAAAGCTTACT 1428
QY 841 ATATCTCCAAATCTCAATTTTCTGGGCAACTCTCTGAGCTATGAGAAAGATTAAAGAC 900
Db 1429 ATATCTCCAAATCTCAATTTTCTGGGCAACTCTCTGAGCTATGAGAAAGATTAAAGAC 1488
QY 901 CAGACTGAGCATCAGAGGCAAGAGCAAACTCAAGCTGCTGACCTGAGAAAGCAAAAT 960
Db 1489 CAGACTGAGCATCAGAGGCAAGAGCAAACTCAAGCTGCTGACCTGAGAAAGCAAAAT 1548
QY 961 GAACTGTCTCTGCTGTCTTCAAGAGGTGAGCAGAAAGGAGAGCGCCCTCACTCAAGCC 1020
Db 1549 GAACTGTCTCTGCTGTCTTCAAGAGGTGAGCAGAAAGGAGAGCGCCCTCACTCAAGCC 1608
QY 1021 TGTGCCGAGCTGTGCTACTCAGAGGAGCAGAGCAAAAGGCGGTGATCCCGCAGCGT 1080
Db 1609 TGTGCCGAGCTGTGCTACTCAGAGGAGCAGAGCAAAAGGCGGTGATCCCGCAGCGT 1668
QY 1081 CCCAGCGTCCAGCGCTGAGCGCTGTGTTAGAGAGCAGCCGCTGTGATCAGCGCTCT 1140
Db 1669 CCCAGCGTCCAGCGCTGAGCGCTGTGTTAGAGAGCAGCCGCTGTGATCAGCGCTCT 1728
QY 1141 AGTGGCTGCACTGTGCGGAGCAGCGCTGAGAGAGCAGCAATTAAGCTCAAGCTTCTTC 1200
Db 1729 AGTGGCTGCACTGTGCGGAGCAGCGCTGAGAGAGCAGCAATTAAGCTCAAGCTTCTTC 1788
QY 1201 TCTCTGATATCAAAATCAGTTTATATTGAGCAGAGATGAGCAGATCTTACATGAGCTTC 1260
Db 1789 TCTCTGATATCAAAATCAGTTTATATTGAGCAGAGATGAGCAGATCTTACATGAGCTTC 1848
QY 1261 TCTCTATCAGAAAGATGCTTTGGAATATCAAACTTTCATCTGAGTGGAGCAAC 1320
Db 1849 TCTCTATCAGAAAGATGCTTTGGAATATCAAACTTTCATCTGAGTGGAGCAAC 1908

QY 1321 AAGCTATGCGAGTCTCCCTGTTACAGAACTATCGAGAGACTCCCGAAACAGCTCT 1380
Db 1909 AAGCTATGCGAGTCTCCCTGTTACAGAACTATCGAGAGACTCCCGAAACAGCTCT 1368
QY 1381 GATPAGAGAGAGCGACATCTCCCAAGAGCTGACAGCCCGACGCTTCAACAGCAG 1440
Db 1969 GATPAGAGAGAGCGACATCTCCCAAGAGCTGACAGCCCGACGCTTCAACAGCAG 2028
QY 1441 AGCAGGATTTGCTGCTGACAGACGAGAGCTGACAGCCCGACGAGTCTCTTTA 1500
Db 2029 AGCAGGATTTGCTGCTGACAGACGAGAGCTGACAGCCCGACGAGTCTCTTTA 2088
QY 1501 TCTTCACATGATGAGAGTGGAGAGCTGAGAGACATATACACAGCTTCTTTTCGCG 1560
Db 2089 TCTTCACATGATGAGAGTGGAGAGCTGAGAGACATATACACAGCTTCTTTTCGCG 2148
QY 1561 CTTTTCACCGACGACGACCTTCAAGAGTCTGCTGCTGCTGCTGCTTAAAGGCTGCGAC 1620
Db 2149 CTTTTCACCGACGACGACCTTCAAGAGTCTGCTGCTGCTGCTTAAAGGCTGCGAC 2208
QY 1621 TCGGATATCTTGGCCCCCGACAGCTCTACCCCTGACGACGAGCTGATTTTGGC 1680
Db 2209 TCGGATATCTTGGCCCCCGACAGCTCTACCCCTGACGACGAGCTGATTTTGGC 2268
QY 1681 ACAAGTCTCTACAGCTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db 2269 ACAAGTCTCTACAGCTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2328
QY 1741 GCTTACAGCTGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 2329 GCTTACAGCTGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2388
QY 1801 CAGAACCCAGTACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 2389 CAGAACCCAGTACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2448
QY 1861 AAGCACTTTAAAGCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 2449 AAGCACTTTAAAGCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2508
QY 1921 AAGTACCGGGAAGAGCTGCGGGAAGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 2509 AAGTACCGGGAAGAGCTGCGGGAAGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2568
QY 1981 ATCATTAGAGTCTCC 1995
Db 2569 ATCATTAGAGTCTCC 2583

RESULT 8
AF506796 3566 bp mRNA linear PRI 30-OCT-2003
LOCUS Homo sapiens dual specificity phosphatase 16 (DUSP16) mRNA,
DEFINITION complete cds.
ACCESSION AF506796
VERSION AF506796.1 GI:25573087
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 3566)
AUTHORS Hoonnaert,I., Marynen,P., Goris,J., Sciort,R. and Baens,M.
TITLE MAPK phosphatase DUSP16/MKP-7, a candidate tumor suppressor for
chromosome region 12p12-13, reduces BCR-ABL-induced transformation
JOURNAL 14586399
PUBMED 2 (bases 1 to 3566)
REFERENCE Hoonnaert,I., Marynen,P. and Baens,M.
AUTHORS Direct Submission
TITLE Submitted (26-APR-2002) Department for Human Genetics-Flanders
JOURNAL Interuniversity Institute for Biotechnology (VIB), Katholieke
Universiteit Leuven, Herestraat 49, Leuven B-3000, Belgium

FEATURES
source location/Qualifiers
1..3566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p13-p12"
/ligase_type="bone marrow"
1..3566
/gene="DUSP16"
/note="synonym: MKP7"
633..2630
/gene="DUSP16"
/BC_number="3.1.3.48"
/BC_number="3.1.3.16"
/codon_start=1
/product="dual specificity phosphatase 16"
/protein_id="AA075120.1"
/db_xref="GI:25573088"
/translation="MAHEMIGTOIVTERLVALLSGREKVLILDSRPVEVNTSHILE
AINNGSKLMRRLOODKVLITELIQSHAKKVIDSGQKVVVDQSDQVAVLSDDC
FLTVLGLKERSFNSVHLIAGFAFSPRCFGLCEGSTLVPCTISQCLPVANIGPT
RILPNLYGCDRDVNLKEMQONGIGVNLASNTCPXPDPIFESHFLRPVNDSECK
ILPMLDKSVDFIEKAKANGCVLHVCLAIGRSATIALAYIMKMDMSLDAYEFKE
KRPTISPVNFILGOLDYERKIKNOTGASGPKSLKULHLEKPMPEVPAYSGOKSE
TPLSPPCDASATSEAAAGORPVHPASVSPVQSLLEDSPLNALSGHLISARLBD
SNLKRSLDIKVSYSASMAASLHGFSSSEDLBYRYKRETTLDGTRKLCQSPVOE
LSBOTPETSPPKEASIPKLIQTARPSDSOSKRLHSVTSSTGTAORSLLSPRSGS
VEDVYHNSFLGLSTSOQLTKSGAGLKGWHSIDILAPOTSPBLTSSWFATPSHR
YSASAIYGGSAVSAYVSCSLPTCGDQVYSVRROKRPDSRADSRSHWEEBPFKOFK
RSCOMERGESEIMSENREBELKVGSGSPSGMEILEVS"

ORIGIN
Query Match 94.9%; Score 1893; DB 9; Length 3566;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATTGGAACCAATTGTACTGAGAGGTTGGTGGCTCTGTGGA 60
Db 633 ATGGCCCATGAGATGATTGGAACCAANTTTTACTGAGAGGTTGGTGGCTCTGTGGA 692
QY 61 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCCGGCATTTGTGAATCAATACATCC 120
Db 693 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCCGGCATTTGTGAATCAATACATCC 752
QY 121 CACATTTTGAAGCATTATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 180
Db 753 CACATTTTGAAGCATTATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 812
QY 181 GACAAAGTGTAAATTACAGAGCTCATCGAGATTCAGCGAAACATAAGTTGACATTGAT 240
Db 813 GACAAAGTGTAAATTACAGAGCTCATCGAGATTCAGCGAAACATAAGTTGACATTGAT 872
QY 241 TGCAGTCAGAAAGTTGATTTAGATCAAAAGCTCCCAAGATGTGCTCTCTTCA 300
Db 873 TGCAGTCAGAAAGTTGATTTAGATCAAAAGCTCCCAAGATGTGCTCTCTTCA 932
QY 301 GACTGTTTTCACATGTAATCTTGGGTAAACTGGAAGAGCTTCAACTCTGTTACCTG 360
Db 933 GACTGTTTTCACATGTAATCTTGGGTAAACTGGAAGAGCTTCAACTCTGTTACCTG 992
QY 361 CTTGCAAGTGGGTTTGTGAGATTCTCGTGTGTTCCCTGGCCCTGTGAAGAAATCC 420
Db 993 CTTGCAAGTGGGTTTGTGAGATTCTCGTGTGTTCCCTGGCCCTGTGAAGAAATCC 1052
QY 421 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGCTTACCTTGCCCAATGGGCCAAC 480
Db 1053 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGCTTACCTTGCCCAATGGGCCAAC 1112
QY 481 CGAATTTTCCCAATCTTTATCTTGGTGCACGAGATGCTCTCAACAGAGCTGATA 540
Db 1113 CGAATTTTCCCAATCTTTATCTTGGTGCACGAGATGCTCTCAACAGAGCTGATA 1172

QY 541 CAGCAGATGGATGGTATGTTAAATGACAGTATACCTGTCCAAAGCCTGACTT 600
| | | | |
Db 1173 CAGCAGATGGATGGTATGTTAAATGACAGTATACCTGTCCAAAGCCTGACTT 1232
| | | | |
QY 601 ATCCCGAGTCTCATTTCTGCGGTGCTGTGATGACAGCTTTGTGAAATTTTG 660
| | | | |
Db 1233 ATCCCGAGTCTCATTTCTGCGGTGCTGTGATGACAGCTTTGTGAAATTTTG 1292
| | | | |
QY 661 CCGTGTGTGACAAATAGTATGATTTGATGAGAAAGCAAAAGCTTCATGATGTT 720
| | | | |
Db 1293 CCGTGTGTGACAAATAGTATGATTTGATGAGAAAGCAAAAGCTTCATGATGTT 1352
| | | | |
QY 721 CTAGTGACTGTTTGTAGTGGATCTCCGCTCCGCGACCATGCTATGCTTATCATG 780
| | | | |
Db 1353 CTAGTGACTGTTTGTAGTGGATCTCCGCTCCGCGACCATGCTTATGCTTATCATG 1412
| | | | |
QY 781 AAGAGATGACATGCTTTAGTGAAGCTTAAGATTTGTGAAGAAAGAAAGCTTACT 840
| | | | |
Db 1413 AAGAGATGACATGCTTTAGTGAAGCTTAAGATTTGTGAAGAAAGAAAGCTTACT 1472
| | | | |
QY 841 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGACTATGAGAAAGATTTAAGAAC 900
| | | | |
Db 1473 ATATCTCCAACTTCAATTTTCTGGGCAACTCTGACTATGAGAAAGATTTAAGAAC 1532
| | | | |
QY 901 CAGACTGGAGCATCAGGGCCAAAGAGCAACTCAAGCTGTGACCTGGAAGCCCAAT 960
| | | | |
Db 1533 CAGACTGGAGCATCAGGGCCAAAGAGCAACTCAAGCTGTGACCTGGAAGCCCAAT 1592
| | | | |
QY 961 GAACTGTCTCTGTCTCAGAGGGTGTGACAGAAAGCAGACGCTTCACTCCACC 1020
| | | | |
Db 1593 GAACTGTCTCTGTCTCAGAGGGTGTGACAGAAAGCAGACGCTTCACTCCACC 1652
| | | | |
QY 1021 TGTGCGCACTGTCTCACTCAGAGGCGACGACAAAGCCCTGCACTCCCGCAGGCTG 1080
| | | | |
Db 1653 TGTGCGCACTGTCTCACTCAGAGGCGACGACAAAGCCCTGCACTCCCGCAGGCTG 1712
| | | | |
QY 1081 CCCAGCGTGGCCAGCGGCGACCGGTGCTGTTAGAGACAGCCCGCTGATCAGGGCTC 1140
| | | | |
Db 1713 CCCAGCGTGGCCAGCGGCGACCGGTGCTGTTAGAGACAGCCCGCTGATCAGGGCTC 1772
| | | | |
QY 1141 AGTGGGTGCACTGTCTCCGACAGAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCCCTG 1200
| | | | |
Db 1773 AGTGGGTGCACTGTCTCCGACAGAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCCCTG 1832
| | | | |
QY 1201 TCTCTGATATCAATATGATTTATATTTCAAGCAGCATGCGACGATCTTACATGCTTC 1260
| | | | |
Db 1833 TCTCTGATATCAATATGATTTATATTTCAAGCAGCATGCGACGATCTTACATGCTTC 1892
| | | | |
QY 1261 TCTCTATCAGAAATGCTTTGGAATCTAATACTTCCACTCTCTGATGAGGACCAAC 1320
| | | | |
Db 1893 TCTCTATCAGAAATGCTTTGGAATCTAATACTTCCACTCTCTGATGAGGACCAAC 1952
| | | | |
QY 1321 AAGCTATGCAAGTTCCTCTGTTCAAGAACTATCGAGAGACCTCCCGAAACCAAGCTCT 1380
| | | | |
Db 1953 AAGCTATGCAAGTTCCTCTGTTCAAGAACTATCGAGAGACCTCCCGAAACCAAGCTCT 2012
| | | | |
QY 1381 GATTAAGAGAGAGCAGCATCTCCCAAGAGAGCTGACAGCCGCGCTTCAAGACGACG 1440
| | | | |
Db 2013 GATTAAGAGAGAGCAGCATCTCCCAAGAGAGCTGACAGCCGCGCTTCAAGACGACG 2072
| | | | |
QY 1441 AGAAGAGATTTGATTCGGTCAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
| | | | |
Db 2073 AGAAGAGATTTGATTCGGTCAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2132
| | | | |
QY 1501 TCTCACTGATCGAATGAG 1560
| | | | |
Db 2133 TCTCACTGATCGAATGAG 2192
| | | | |
QY 1561 CTTTCCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1620
| | | | |
Db 2193 CTTTCCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2252
| | | | |
QY 1621 TCGGATATCTTTGGCCCCCAGACCTCTAACCCCTTCCCTGACAGCAGCTGTATTTTGGC 1680
| | | | |

Db 2253 TCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGTATTTTGGC 2312
| | | | |
QY 1681 ACAGAGTCTCTACACTTCTACTGCTTCAAGCATTTAGAGAGAGAGAGAGAGAGAGAG 1740
| | | | |
Db 2313 ACAGAGTCTCTACACTTCTACTGCTTCAAGCATTTAGAGAGAGAGAGAGAGAGAGAG 2372
| | | | |
QY 1741 GCTTACAGTGCAG 1800
| | | | |
Db 2373 GCTTACAGTGCAG 2432
| | | | |
QY 1801 CAGAGCCAAAGTGCAG 1860
| | | | |
Db 2433 CAGAGCCAAAGTGCAG 2492
| | | | |
QY 1861 AAGCAGTTTAAACGAG 1920
| | | | |
Db 2493 AAGCAGTTTAAACGAG 2552
| | | | |
QY 1921 AGGTCAAGGAG 1980
| | | | |
Db 2553 AGGTCAAGGAG 2612
| | | | |
QY 1981 ATCATGAGTCTCC 1995
| | | | |
Db 2613 ATCATGAGTCTCC 2627
| | | | |
RESULT 9
AX374994 3766 bp DNA linear PAT 01-MAR-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Tang,Y.T., Elliott,V.S., Rankumar,J., Yao,M.G., Burford,N.,
Wang,Y.E., Stewart,E.A., Gandhi,A.R., Patterson,C., Lee,B.A.,
Hafalla,A.J., Lu,D.A., Tribouley,C.M., Griffin,J.A., Baughn,M.R.,
Yue,H., Warren,B.A., Nguyen,D.B. and Walla,N.K.
Protein phosphatases
Patent: WO 0210363-A 17 07-FEB-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source
1..3766
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7480570CB1"
ORIGIN
Query Match 94.9%; Score 1893; DB 6; Length 3766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCATGAGATGATGGAATCTCAATTTGTAAGAGAGTGTGGCTGTGGAA 60
| | | | |
Db 538 ATGGCCATGAGATGATGGAATCTCAATTTGTAAGAGAGTGTGGCTGTGGAA 597
| | | | |
QY 61 AGTGAACGGAAGAAAGTGTCTAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
| | | | |
Db 598 AGTGAACGGAAGAAAGTGTCTAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 657
| | | | |
QY 121 CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATAGAGAGAGAGAGAGAGAGAG 180
| | | | |
Db 658 CACATTTTGAAGCCATTATATCACTGCTCAAGCTTATAGAGAGAGAGAGAGAGAGAG 717
| | | | |
QY 181 GACAAAGTGTATTTAGAGAGCTCATCCAGCATTTGAGGAAACATTAAGTTGATTTGAT 240
| | | | |

Db 718 GACAAAGTGAATTAACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 777
Qy 241 TGCAGTGAAGAGTGTAGTTAGATTAAGTCCCAAGTGTGCTCTCTCTTCA 300
Db 778 TGCAGTGAAGAGTGTAGTTAGATTAAGTCCCAAGTGTGCTCTCTCTTCA 837
Qy 301 GACTGTTTCTCAGTACTTCTGGGTAACTGAGAAAGAGCTTCACTGTTTCACTG 360
Db 838 GACTGTTTCTCAGTACTTCTGGGTAACTGAGAAAGAGCTTCACTGTTTCACTG 897
Qy 361 CTTCAGAGTGGGTTTGTCTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGAAAATCC 420
Db 898 CTTCAGAGTGGGTTTGTCTGAGTTCTCTGTTGTTTCCCTGAGCTCTGTGAAGAAAATCC 957
Qy 421 ACTTAAGTCCCTACCTGATTTCTCAGCTTGTCTTACTTGTGCGAACATTGGGCCAAC 480
Db 958 ACTTAAGTCCCTACCTGATTTCTCAGCTTGTCTTACTTGTGCGAACATTGGGCCAAC 1017
Qy 481 CGAATTCCTCCCAATCTTTATCTTGGCTGCAGAGATGTCTCAACAGAGCTGATA 540
Db 1018 CGAATTCCTCCCAATCTTTATCTTGGCTGCAGAGATGTCTCTCAACAGAGCTGATG 1077
Qy 541 CAGAGATAGAGTGTGTTATGTGTAAATGCCAGCTATACTGTCCAAAGCCTGAATT 600
Db 1078 CAGAGATAGAGTGTGTTATGTGTAAATGCCAGCAATCTGTCCAAAGCCTGAATT 1137
Qy 601 ATCCCGAGTCTCATTTCTGCGGTGCTGTGAATGACGCTTTGTGAGAAATTTTG 660
Db 1138 ATCCCGAGTCTCATTTCTGCGGTGCTGTGAATGACGCTTTGTGAGAAATTTTG 1197
Qy 661 CCGTGTGTGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCTTCAATGATGTGT 720
Db 1198 CCGTGTGTGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCTTCAATGATGTGT 1257
Qy 721 CTAGTGAACGTTTATGATGGGATCTCCCGTCCGCCACATGCTATTCGCTATCATCATG 780
Db 1258 CTAGTGAACGTTTATGATGGGATCTCCCGTCCGCCACATGCTATTCGCTATCATCATG 1317
Qy 781 AAGAGATGACATGTCCTTAGATGAAGCTTACAGATTTGTGAAGAAAAGAAAGCACTACT 840
Db 1318 AAGAGATGACATGTCCTTAGATGAAGCTTACAGATTTGTGAAGAAAAGAAAGCACTACT 1377
Qy 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTTAAGAAC 900
Db 1378 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTTAAGAAC 1437
Qy 901 CAGACTGAGGACATCAGGGCCAAAGACAACTAACCTGTGACCTTGAAGAACCAAT 960
Db 1438 CAGACTGAGGACATCAGGGCCAAAGACAACTAACCTGTGACCTTGAAGAACCAAT 1497
Qy 961 GAACCTGTCCCTGCTCTCAGAGGTTGACAGAAAGGAGAGCGCCCTCAGTCCACCC 1020
Db 1498 GAACCTGTCCCTGCTCTCAGAGGTTGACAGAAAGGAGAGCGCCCTCAGTCCACCC 1557
Qy 1021 TGTGCCGACTGTGCTACTTCAAGGAGCAGAGCAAAAGGCCGTGATCCCGCAGCGTG 1080
Db 1558 TGTGCCGACTGTGCTACTTCAAGGAGCAGAGCAAAAGGCCGTGATCCCGCAGCGTG 1617
Qy 1081 CCAGGCTGCCAGCGGTGACGCGTGTGTTAGAGAACAGCCCGTGTGTAAGCGCGCTC 1140
Db 1618 CCAGGCTGCCAGCGGTGACGCGTGTGTTAGAGAACAGCCCGTGTGTAAGCGCGCTC 1677
Qy 1141 AGTGGGCTGCACCTGTCCGACAGAGGCTGAGAAAGAGCAATTAAGTCAAGGCTTCTTC 1200
Db 1678 AGTGGGCTGCACCTGTCCGACAGAGGCTGAGAAAGAGCAATTAAGTCAAGGCTTCTTC 1737
Qy 1201 TCTCTGATATCAATCAAGTTTCAATTCAGCCAGCATGCGCATCTTACATGCTTC 1260
Db 1738 TCTCTGATATCAATCAAGTTTCAATTCAGCCAGCATGCGCATCTTACATGCTTC 1797
Qy 1261 TCCCTATCAGAGATGCTTTGAATACTAACAACTTCCACTCTGTGATGGACCAAC 1320
Db 1798 TCCCTATCAGAGATGCTTTGAATACTAACAACTTCCACTCTGTGATGGACCAAC 1857

Qy 1321 AAGCTATGCCAGTTCCTCCCTGTTCAGGAATATCGAGACAGATCCCGAAACAGTCTCT 1380
Db 1858 AAGCTATGCCAGTTCCTCCCTGTTCAGGAATATCGAGACAGATCCCGAAACAGTCTCT 1917
Qy 1381 GATTAAGAGAAAGCCAGCATTCCTCCAAAGAAAGCTGACACCGCCAGGCTTCAACAGCCAG 1440
Db 1918 GATTAAGAGAAAGCCAGCATTCCTCCAAAGAAAGCTGACACCGCCAGGCTTCAACAGCCAG 1977
Qy 1441 AGCAAGCATTCATTCGCTGAGAACCCAGACAGTGGACCGCCAGAGGCTCCCTTTTA 1500
Db 1978 AGCAAGCATTCATTCGCTGAGAACCCAGACAGTGGACCGCCAGAGGCTCCCTTTTA 2037
Qy 1501 TCTCCATGTCATCGAAGTGGAGCGCTGAGAGCAATTACACACAGCATCTCTTTTGGGC 1560
Db 2038 TCTCCATGTCATCGAAGTGGAGCGCTGAGAGCAATTACACACAGCATCTCTTTTGGGC 2097
Qy 1561 CTTTCCACAGCAGACAGACCTCAAGAGTGTGCTGGCTTGAAGGCTTGGGCAC 1620
Db 2098 CTTTCCACAGCAGACAGACCTCAAGAGTGTGCTGGCTTGAAGGCTTGGGCAC 2157
Qy 1621 TCGAATATCTTGGCCCCCAGACCTCTACCCCTTCCAGACAGAGCTGTGATTTTGGC 1680
Db 2158 TCGAATATCTTGGCCCCCAGACCTCTACCCCTTCCAGACAGAGCTGTGATTTTGGC 2217
Qy 1681 ACAGAGTCTCTACATTTCTACTTGTGCTGACGATCTACAGAGGAGTGGCCAGTTACTCT 1740
Db 2218 ACAGAGTCTCTACATTTCTACTTGTGCTGACGATCTACAGAGGAGTGGCCAGTTACTCT 2277
Qy 1741 GCTTACAGCTGAGCAGCTGCCCCACTTGGGGAGACCAAGCTATTCGTGTCGAGGCGG 1800
Db 2278 GCTTACAGCTGAGCAGCTGCCCCACTTGGGGAGACCAAGCTATTCGTGTCGAGGCGG 2337
Qy 1801 CAGAAAGCAAGTGAAGAGCTGACTCGCGGCGAGAGCTGGACATGAAGAGGCCCTTTGAA 1860
Db 2338 CAGAAAGCAAGTGAAGAGCTGACTCGCGGCGAGAGCTGGACATGAAGAGGCCCTTTGAA 2397
Qy 1861 AAGCATTTAAACGAGAAAGCTGCGCAATTTGAGAGAGCTATGTCAGAGAAC 1920
Db 2398 AAGCATTTAAACGAGAAAGCTGCGCAATTTGAGAGAGCTATGTCAGAGAAC 2457
Qy 1921 AGGTACGAGGAGAGCTGGGAAAGTGGGAGTCACTTCTTGGGAGAGATGGA 1980
Db 2458 AGGTACGAGGAGAGCTGGGAAAGTGGGAGTCACTTCTTGGGAGAGATGGA 2517
Qy 1981 ATCATTGAGTCTCC 1995
Db 2518 ATCATTGAGTCTCC 2532

RESULT 10
BD171157
LOCUS BD171157 4790 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel gene and protein encoded thereby.
ACCESSION BD171157.1 GI:27876969
VERSION BD171157.1 04-JUL-2002
KEYWORDS WO 02052005-A/13.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 4790)
AUTHORS Ohara,O., Nagase,T. and Nakajima,D.
TITLE Novel gene and protein encoded thereby
JOURNAL Patent: WO 02052005-A 13 04-JUL-2002;
KAZUSA DNA RESEARCH INSTITUTE,OSAMU OHARA,TAKAHIRO NAGASE, DAISUKE
NAKAJIMA
OS Homo sapiens (human)
PN WO 02052005-A/13
PD 04-JUL-2002
PF 20-DEC-2001 WO 2001JP011217
PI 22-DEC-2000 JP 00P 389742
PI OSAMU OHARA,TAKAHIRO NAGASE,DAISUKE NAKAJIMA

PC C12N15/12, C07K14/47
CC Novel gene and protein encoded thereby
FH Key Location/Qualifiers
FT CDS (184)..(2178).

FEATURES
Source 1..4790
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 94.9%; Score 1893; DB 6; Length 4790;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCCCAGATGATGATGGAACCTCAATTTGTAATGAGAGGTTGGTCTGCTGGAA 60
DB ATGCCCAGATGATGATGGAACCTCAATTTGTAATGAGAGGTTGGTCTGCTGGAA 184
QY 61 AGTGAACCGAAGAGCTGCTAAATGATAGCCGCGCATTTGGAAATACATACATCC 120
DB AGTGAACCGAAGAGCTGCTAAATGATAGCCGCGCATTTGGAAATACATACATCC 244
QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 180
DB CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 304
QY 181 GACAAAGTGTAAATTCAGAGCTCATCAGCAATTCAGCGAACAATAAGTTGACATTTAT 240
DB GACAAAGTGTAAATTCAGAGCTCATCAGCAATTCAGCGAACAATAAGTTGACATTTAT 364
QY 241 TGCAGTCAGAAAGTTGATGATTAAGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 300
DB TGCAGTCAGAAAGTTGATGATTAAGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 424
QY 424 TGCAGTCAGAAAGTTGATGATTAAGATCAAGCTCCCAAGATGTTGCTCTCTCTCA 483
QY 301 GACTGTTTCTCACTGACTTCTGGGTAATGGAAGAGCTTCAACTCTGTTCACTG 360
DB GACTGTTTCTCACTGACTTCTGGGTAATGGAAGAGCTTCAACTCTGTTCACTG 484
QY 484 GACTGTTTCTCACTGACTTCTGGGTAATGGAAGAGCTTCAACTCTGTTCACTG 543
DB GACTGTTTCTCACTGACTTCTGGGTAATGGAAGAGCTTCAACTCTGTTCACTG 544
QY 544 CTTCAGAGTGGGTTGCTGAGTTCTCTGTTGTTCCCTGGCCCTGTGAAAGAAATCC 603
DB CTTCAGAGTGGGTTGCTGAGTTCTCTGTTGTTCCCTGGCCCTGTGAAAGAAATCC 603
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGTCTTACCTGTTGCAACATTTGGCCCAAC 480
DB ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGTCTTACCTGTTGCAACATTTGGCCCAAC 604
QY 604 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGTCTTACCTGTTGCAACATTTGGCCCAAC 663
DB ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGTCTTACCTGTTGCAACATTTGGCCCAAC 481
QY 481 CGAATTCCTCCCAATCTTTATCTTGGCTGCGCAGGAGATGCTCTCAACAAGAGCTGATA 540
DB CGAATTCCTCCCAATCTTTATCTTGGCTGCGCAGGAGATGCTCTCAACAAGAGCTGATA 664
QY 664 CGAATTCCTCCCAATCTTTATCTTGGCTGCGCAGGAGATGCTCTCAACAAGAGCTGATA 723
DB CGAATTCCTCCCAATCTTTATCTTGGCTGCGCAGGAGATGCTCTCAACAAGAGCTGATA 541
QY 541 CAGCAGAAATGGATTTGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 600
DB CAGCAGAAATGGATTTGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 724
QY 724 CAGCAGAAATGGATTTGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 783
DB CAGCAGAAATGGATTTGTTATGTTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT 601
QY 601 ATCCCGAGTCTCATTTCTCGCGTGTGCTGTGAATGACAGCTTTTGTGAATAATTTTG 660
DB ATCCCGAGTCTCATTTCTCGCGTGTGCTGTGAATGACAGCTTTTGTGAATAATTTTG 784
QY 784 ATCCCGAGTCTCATTTCTCGCGTGTGCTGTGAATGACAGCTTTTGTGAATAATTTTG 843
DB ATCCCGAGTCTCATTTCTCGCGTGTGCTGTGAATGACAGCTTTTGTGAATAATTTTG 661
QY 661 CCGTGGTTGACAAATCAGTAGATTTTCAATTGAGAAAGCAAAAGCCTCCCAATGATGTT 720
DB CCGTGGTTGACAAATCAGTAGATTTTCAATTGAGAAAGCAAAAGCCTCCCAATGATGTT 844
QY 844 CCGTGGTTGACAAATCAGTAGATTTTCAATTGAGAAAGCAAAAGCCTCCCAATGATGTT 903
DB CCGTGGTTGACAAATCAGTAGATTTTCAATTGAGAAAGCAAAAGCCTCCCAATGATGTT 721
QY 721 CTAGTGCACTGTTAGCTGGGATCTCCGCTCGGCCACATCGCTATCGCTACATCATG 780
DB CTAGTGCACTGTTAGCTGGGATCTCCGCTCGGCCACATCGCTATCGCTACATCATG 904
QY 904 CTAGTGCACTGTTAGCTGGGATCTCCGCTCGGCCACATCGCTATCGCTACATCATG 963
DB CTAGTGCACTGTTAGCTGGGATCTCCGCTCGGCCACATCGCTATCGCTACATCATG 781
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTTGTGAAGAAAAAGACCTACT 840
DB AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTTGTGAAGAAAAAGACCTACT 964
QY 964 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTTGTGAAGAAAAAGACCTACT 1023
DB AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTTGTGAAGAAAAAGACCTACT 841
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCCTGAGCTATGAGAAAGATTGAAGAC 900
DB ATATCTCAAACTTCAATTTTCTGGGCAACTCCTGAGCTATGAGAAAGATTGAAGAC 900

DB 1024 ATATCTCAAACTTCAATTTTCTGGGCAACTCCTGAGCTATGAGAAAGATTGAAGAC 1083
QY 901 CAGACTGGAGCATCAGGGCCAAAGCAAACTCAAGCTCTGCACTCTGGAGAACCAAT 960
DB CAGACTGGAGCATCAGGGCCAAAGCAAACTCAAGCTCTGCACTCTGGAGAACCAAT 1084
QY 961 GAACTGTCTCCCTGCTCTCAGAGGGTGAACAGAAAGAGAGAGCGCCCTCAGTCCACC 1020
DB GAACTGTCTCCCTGCTCTCAGAGGGTGAACAGAAAGAGAGAGCGCCCTCAGTCCACC 1144
QY 1144 GAACTGTCTCCCTGCTCTCAGAGGGTGAACAGAAAGAGAGAGCGCCCTCAGTCCACC 1203
DB GAACTGTCTCCCTGCTCTCAGAGGGTGAACAGAAAGAGAGAGCGCCCTCAGTCCACC 1021
QY 1021 TGTGCGCACTGCTCACTCAGAGGAGAGAGCAAAAGCCGTGATCCCGCAGCTG 1080
DB TGTGCGCACTGCTCACTCAGAGGAGAGAGCAAAAGCCGTGATCCCGCAGCTG 1204
QY 1204 TGTGCGCACTGCTCACTCAGAGGAGAGAGCAAAAGCCGTGATCCCGCAGCTG 1263
DB TGTGCGCACTGCTCACTCAGAGGAGAGAGCAAAAGCCGTGATCCCGCAGCTG 1081
QY 1081 CCCAGGCTGCCAGGCTGAGGCTGCTGTTAGAGAGACAGCCCGCTGGTACAGGGCTC 1140
DB CCCAGGCTGCCAGGCTGAGGCTGCTGTTAGAGAGACAGCCCGCTGGTACAGGGCTC 1264
QY 1264 CCCAGGCTGCCAGGCTGAGGCTGCTGTTAGAGAGACAGCCCGCTGGTACAGGGCTC 1323
DB CCCAGGCTGCCAGGCTGAGGCTGCTGTTAGAGAGACAGCCCGCTGGTACAGGGCTC 1141
QY 1141 AGTGGCTGACACTGTCGCGCAGACAGGCTGAAAGCAGCAATTAAGCTCAAGCGTCTTC 1200
DB AGTGGCTGACACTGTCGCGCAGACAGGCTGAAAGCAGCAATTAAGCTCAAGCGTCTTC 1324
QY 1324 AGTGGCTGACACTGTCGCGCAGACAGGCTGAAAGCAGCAATTAAGCTCAAGCGTCTTC 1383
DB AGTGGCTGACACTGTCGCGCAGACAGGCTGAAAGCAGCAATTAAGCTCAAGCGTCTTC 1201
QY 1201 TCTCTGATATCAAAATCAGTTTCAATTAATTCAGCAGCATGAGCATCTTACATGCTTC 1260
DB TCTCTGATATCAAAATCAGTTTCAATTAATTCAGCAGCATGAGCATCTTACATGCTTC 1384
QY 1384 TCTCTGATATCAAAATCAGTTTCAATTAATTCAGCAGCATGAGCATCTTACATGCTTC 1443
DB TCTCTGATATCAAAATCAGTTTCAATTAATTCAGCAGCATGAGCATCTTACATGCTTC 1261
QY 1261 TCTCTATCAAGAGATCTTTGGAAATCTAACAACTTCCACTTCTGATGGGACCAAC 1320
DB TCTCTATCAAGAGATCTTTGGAAATCTAACAACTTCCACTTCTGATGGGACCAAC 1444
QY 1444 TCTCTATCAAGAGATCTTTGGAAATCTAACAACTTCCACTTCTGATGGGACCAAC 1503
DB TCTCTATCAAGAGATCTTTGGAAATCTAACAACTTCCACTTCTGATGGGACCAAC 1321
QY 1321 AAGCTATGCGAGTTCTCCCTGTTGAGAACTATGGAGACAGCTCCGAAACCACTGCT 1380
DB AAGCTATGCGAGTTCTCCCTGTTGAGAACTATGGAGACAGCTCCGAAACCACTGCT 1504
QY 1504 AAGCTATGCGAGTTCTCCCTGTTGAGAACTATGGAGACAGCTCCGAAACCACTGCT 1563
DB AAGCTATGCGAGTTCTCCCTGTTGAGAACTATGGAGACAGCTCCGAAACCACTGCT 1381
QY 1381 GATTAAGAGAAAGCCAGCATCTCCCAAGAGCTGCAAGCCGACCTTCAAGACCAAG 1440
DB GATTAAGAGAAAGCCAGCATCTCCCAAGAGCTGCAAGCCGACCTTCAAGACCAAG 1564
QY 1564 GATTAAGAGAAAGCCAGCATCTCCCAAGAGCTGCAAGCCGACCTTCAAGACCAAG 1623
DB GATTAAGAGAAAGCCAGCATCTCCCAAGAGCTGCAAGCCGACCTTCAAGACCAAG 1441
QY 1441 AGCAAGCATTTGATTCGCTCAGAACACAGACAGTGGCAACCGCCAGAGCTCTTTTA 1500
DB AGCAAGCATTTGATTCGCTCAGAACACAGACAGTGGCAACCGCCAGAGCTCTTTTA 1624
QY 1624 AGCAAGCATTTGATTCGCTCAGAACACAGACAGTGGCAACCGCCAGAGCTCTTTTA 1683
DB AGCAAGCATTTGATTCGCTCAGAACACAGACAGTGGCAACCGCCAGAGCTCTTTTA 1501
QY 1501 TCTCCACGTCATTCGAAGTGGAGCGTGGAGGCAATTAACACAGCTCTCTTCCG 1560
DB TCTCCACGTCATTCGAAGTGGAGCGTGGAGGCAATTAACACAGCTCTCTTCCG 1684
QY 1684 TCTCCACGTCATTCGAAGTGGAGCGTGGAGGCAATTAACACAGCTCTCTTCCG 1743
DB TCTCCACGTCATTCGAAGTGGAGCGTGGAGGCAATTAACACAGCTCTCTTCCG 1561
QY 1561 CTTTCACAGCAGCAGACGACCTCAAGAGTCTGCTGGCTGGGCTTAAAGGCTGGAC 1620
DB CTTTCACAGCAGCAGACGACCTCAAGAGTCTGCTGGCTGGGCTTAAAGGCTGGAC 1744
QY 1744 CTTTCACAGCAGCAGACGACCTCAAGAGTCTGCTGGCTGGGCTTAAAGGCTGGAC 1803
DB CTTTCACAGCAGCAGACGACCTCAAGAGTCTGCTGGCTGGGCTTAAAGGCTGGAC 1621
QY 1621 TCGGATATCTTGGCCGCCAGACCTCTACCCCTTCCCTGACACAGAGCTGTATTTTGGC 1680
DB TCGGATATCTTGGCCGCCAGACCTCTACCCCTTCCCTGACACAGAGCTGTATTTTGGC 1804
QY 1804 TCGGATATCTTGGCCGCCAGACCTCTACCCCTTCCCTGACACAGAGCTGTATTTTGGC 1863
DB TCGGATATCTTGGCCGCCAGACCTCTACCCCTTCCCTGACACAGAGCTGTATTTTGGC 1681
QY 1681 ACAGAGTCTCTACACTTCTTACTCTGCTCAGACCATTAAGAGGAGCTGCCAGTTACTCT 1740
DB ACAGAGTCTCTACACTTCTTACTCTGCTCAGACCATTAAGAGGAGCTGCCAGTTACTCT 1864
QY 1864 ACAGAGTCTCTACACTTCTTACTCTGCTCAGACCATTAAGAGGAGCTGCCAGTTACTCT 1923
DB ACAGAGTCTCTACACTTCTTACTCTGCTCAGACCATTAAGAGGAGCTGCCAGTTACTCT 1741
QY 1741 GCCTACAGCTGACAGCAGCTGCGCACTTGCAGAGCAAGTCTATCTGTGCGCAGCG 1800
DB GCCTACAGCTGACAGCAGCTGCGCACTTGCAGAGCAAGTCTATCTGTGCGCAGCG 1924
QY 1924 GCCTACAGCTGACAGCAGCTGCGCACTTGCAGAGCAAGTCTATCTGTGCGCAGCG 1983
DB GCCTACAGCTGACAGCAGCTGCGCACTTGCAGAGCAAGTCTATCTGTGCGCAGCG 1801
QY 1801 CAGAAAGCAAGTACAGAGCTGACTGCGGCGGAGCTGACATGAAGAGCCCTTTTGA 1860
DB CAGAAAGCAAGTACAGAGCTGACTGCGGCGGAGCTGACATGAAGAGCCCTTTTGA 1984
QY 1984 CAGAAAGCAAGTACAGAGCTGACTGCGGCGGAGCTGACATGAAGAGCCCTTTTGA 2043
DB CAGAAAGCAAGTACAGAGCTGACTGCGGCGGAGCTGACATGAAGAGCCCTTTTGA 1861
QY 1861 AAGCATTTTAAAGCAGAGCTGCCAAATGGAATTTTGAAGAGCATCATGTCAGAGAAC 1920
DB AAGCATTTTAAAGCAGAGCTGCCAAATGGAATTTTGAAGAGCATCATGTCAGAGAAC 2044
QY 2044 AAGCATTTTAAAGCAGAGCTGCCAAATGGAATTTTGAAGAGCATCATGTCAGAGAAC 2103
DB AAGCATTTTAAAGCAGAGCTGCCAAATGGAATTTTGAAGAGCATCATGTCAGAGAAC 1921
QY 1921 AAGTCAAGGGAAGAGCTGGGGAAGAGTGGCAAGTCAAGCTTTTGGGCGAGCATGAA 1980
DB AAGTCAAGGGAAGAGCTGGGGAAGAGTGGCAAGTCAAGCTTTTGGGCGAGCATGAA 2104
QY 2104 AAGTCAAGGGAAGAGCTGGGGAAGAGTGGCAAGTCAAGCTTTTGGGCGAGCATGAA 2163
DB AAGTCAAGGGAAGAGCTGGGGAAGAGTGGCAAGTCAAGCTTTTGGGCGAGCATGAA

Qy	1981	ATCATTGAGGTCCTCC	1995
Db	2164	ATCATTGAGGTCCTCC	2178

	BD183422	BD183422.1	GI:31875622	UJ 2002345492-A/135.	Homo sapiens (human)	Homo sapiens
RESULT 11						
LOCUS	BD183422					
DEFINITION	Novel genes and proteins encoded by the genes.					
ACCESSION	BD183422					
VERSION	BD183422.1					
KEYWORDS	UJ 2002345492-A/135.					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 4790)
Ohara, O., Nagase, T. and Nakajima, D.
Novel genes and proteins encoded by the genes
Patent: JP 2002345492-A 135 03-DEC-2002;

COMMENT OS Homo sapiens (human)

PD	03-DEC-2002
PE	26-FEB-2002 JP 2002049009
PI	OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC	C12N15/09,C07K14/47//A61K31/711,A61K38/00,A61K48/00,A61P25/00,
PC	A61P25/14'
PC	A61P25/18,A61P25/00,C12N15/00,A61K37/02
CC	Novel genes and proteins encoded by the genes FH
Location/Qualifiers	(184) . . (2178) .
FT	CDS

FEATURES	Location/Qualifiers
source	1. .4790

ORIGIN

Query Match	94.9%	Score 1933	DB 6	Length 4790
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1993, Conservative	0	Mismatches	2	Indels 0
				Gaps 0

QY	1	ATGGCCCATGAGATGATTGGAACTCAATTGTATCTGAGAGTGGTGGCTCTGCTGGAA	60
Db	184	ATGGCCCATGAGATGATTGGAACTCAATTGTATCTGAGAGTGGTGGCTCTGCTGGAA	243
QY	61	AGTGAACGGAAAAAGTGCCTCTAATTGATATGCGGGCCATTGTGGAAATCAATCAATCC	120
Db	244	AGTGAACGGAAAAAGTGCCTCTAATTGATATGCGGGCCATTGTGGAAATCAATCAATCC	303
QY	121	CACATTTTGGAAAGCCATTATATCAACTGCTCCAAAGCTTATGAAAGCAAGTTGCAAG	180
Db	304	CACATTTTGGAAAGCCATTATATCAACTGCTCCAAAGCTTATGAAAGCAAGTTGCAAG	363
QY	181	GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTGACATTGAT	240
Db	364	GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTGACATTGAT	423
QY	241	TGCAGTCAGAAAGTTGTAGTTTACGATCAAAAGCTCCAAAGTGTGCTCTCTCTCTTCA	300
Db	424	TGCAGTCAGAAAGTTGTAGTTTACGATCAAAAGCTCCAAAGTGTGCTCTCTCTCTTCA	483
QY	301	GACTGTTTTCTCACTGTACTTCTGGGTAACTGGAAGAGACCTTCAACTGTTCACCTG	360
Db	484	GACTGTTTTCTCACTGTACTTCTGGGTAACTGGAAGAGACCTTCAACTGTTCACCTG	543
QY	361	CTTGCAAGTGGGTTTGTCTGAGTTCTCTGTGTTGTTCCCTGGACCTCTGTGAAGAAATCC	420
Db	544	CTTGCAAGTGGGTTTGTCTGAGTTCTCTGTGTTGTTCCCTGGACCTCTGTGAAGAAATCC	603
QY	421	ACTCTAAGTCCCTACCTGCAATTTCTCAGCCTTGCTTAACTGTGGCAACATTTGGGCCAAC	480

Db	604	ACTCTAGTCCCTACCTGCAATTTCTCAGGCTTGCTTACCTGTTGCCAACTTGGGCCAACCC	663
Qy	481	CGAAATCTTCCCAATCTTATCTTGGCTGCACGACGAGATGTCTCAACAAGAGCTGATA	540
Db	664	CGAAATCTTCCCAATCTTATCTTGGCTGCACGACGAGATGTCTCAACAAGAGAGCTGATG	723
Qy	541	CAGCAGAAATGGGATTTGGTTATGTATTAATGCACTATCCTGTCCAAAGCTTGACTTT	600
Db	724	CAGCAGAAATGGGATTTGGTTATGTATTAATGCACTATCCTGTCCAAAGCTTGACTTT	783
Qy	601	ATCCCCAGATCTCATTTCTCGTGCTGGCTGGAAATGACAGCTTTTGGAGAAAATTTTG	660
Db	784	ATCCCCAGATCTCATTTCTCGTGCTGGCTGGAAATGACAGCTTTTGGAGAAAATTTTG	843
Qy	661	CCGTGGTTGACAAATCAGTAGATTCTTATGAGAAAGAAAAGCTCCAAATGATGTGT	720
Db	844	CCGTGGTTGACAAATCAGTAGATTCTTATGAGAAAGAAAAGCTCCAAATGATGTGT	903
Qy	721	CTAGTGCACTGTTTAGCTGGGATCTCCGCGTCCGCACATCGCTATGCTCAATCATG	780
Db	904	CTAGTGCACTGTTTAGCTGGGATCTCCGCGTCCGCACATCGCTATGCTCAATCATG	963
Qy	781	AAGAGAGATGACATGCTCTTATGATGACCTTACAGATTTGTGAAAAGAAAAGACTTACT	840
Db	964	AAGAGAGATGACATGCTCTTATGATGACCTTACAGATTTGTGAAAAGAAAAGACTTACT	1023
Qy	841	ATATCTCCAAATCTCAATTTTCTGGGCCAACTCTGCACTTATGAGAAAGATTTAAGAC	900
Db	1024	ATATCTCCAAATCTCAATTTTCTGGGCCAACTCTGCACTTATGAGAAAGATTTAAGAC	1083
Qy	901	CAGACTGAGCATCAGGGCCAAAGACAACTCAAGCTGTGCACCTTGAGAGAGCCAAAT	960
Db	1084	CAGACTGAGCATCAGGGCCAAAGACAACTCAAGCTGTGCACCTTGAGAGAGCCAAAT	1143
Qy	961	GAACTGTGCTGCTGTCTCAGAGGCTGGAACAGAAAGCCGACCTCATGTCCACCC	1022
Db	1144	GAACTGTGCTGCTGTCTCAGAGGCTGGAACAGAAAGCCGACCTCATGTCCACCC	1203
Qy	1021	TGTGCCGCACTCTGCTACCTCAGAGGACAGAGCAAAAGGCCGTGTGCATCCGCGACGCTG	1080
Db	1204	TGTGCCGCACTCTGCTACCTCAGAGGACAGAGCAAAAGGCCGTGTGCATCCGCGACGCTG	1263
Qy	1081	CCAGACGCTGCCAGCGTGACGCGCTGTTAGAGGACAGCCCGCTGTGTACAGGCGCTC	1140
Db	1264	CCAGACGCTGCCAGCGTGACGCGCTGTTAGAGGACAGCCCGCTGTGTACAGGCGCTC	1323
Qy	1141	AGTGGGCTGCACCTGTCCGACAGACAGGCTGGAAGACAGCAATAAGCTTAAGCTTCTTC	1200
Db	1324	AGTGGGCTGCACCTGTCCGACAGACAGGCTGGAAGACAGCAATAAGCTTCTTC	1383
Qy	1201	TCTCTGATATCAATCAGTTTCAATTTACACGACGATGGAGAGATCCTTACATGGCTTC	1260
Db	1384	TCTCTGATATCAATCAGTTTCAATTTACACGACGATGGAGAGATCCTTACATGGCTTC	1443
Qy	1261	TCTCTATCAGAAATGCTTTGGAAATATACAAACCTTCACTACTCTGTGATGGGACCAAC	1320
Db	1444	TCTCTATCAGAAATGCTTTGGAAATATACAAACCTTCACTACTCTGTGATGGGACCAAC	1503
Qy	1321	AAGCTATGCGCAATTTCTCCCTGTTCAGAGAACTATCGAGACAGATCCCGAAACCAATCTCT	1380
Db	1504	AAGCTATGCGCAATTTCTCCCTGTTCAGAGAACTATCGAGACAGATCCCGAAACCAATCTCT	1563
Qy	1381	GATTAAGAGGAAGCAGACATCCCAAGAGGTGACACCGCAGGCGCTTACAGACAGCGAG	1440
Db	1564	GATTAAGAGGAAGCAGACATCCCAAGAGGTGACACCGCAGGCGCTTACAGACAGCGAG	1623
Qy	1441	AGCAAGGCAATTCATTTGGTTCAGAAACAGACAGCAAGTGGCACCGGCCAGAGGCTCCTTTTA	1500
Db	1624	AGCAAGGCAATTCATTTGGTTCAGAAACAGACAGCAAGTGGCACCGGCCAGAGGCTCCTTTTA	1663
Qy	1501	TCTTCACTGCATCGAATGGGAGCGCTGAGAGACAAATTACACACCAAGCTTCTTTTTCGGC	1560

Db 1684 TCTCCACTGCATCGAAGTGGGAGCGTGGAGAGCAATTACACACAGCTTCTTTTCGGC 1743

Qy 1561 CTTTCCACCCAGCAGCAGCCTTCAGGAAGTCTGCTGGCTTGAAGGGCTGGAC 1620

Db 1744 CTTTCCACCCAGCAGCAGCCTTCAGGAAGTCTGCTGGCTTGAAGGGCTGGAC 1803

Qy 1621 TCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGTATTTTGGC 1680

Db 1804 TCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGTATTTTGGC 1863

Qy 1681 ACAGAGTCTTCACACTTCTACTCTGCTCAGCAGCATTCAGGAGGAGCTGGCACTTCTCT 1740

Db 1864 ACAGAGTCTTCACACTTCTACTCTGCTCAGCAGCATTCAGGAGGAGCTGGCACTTCTCT 1923

Qy 1741 GCTTACAGCTGCGAGCAGCTGCCCACTTGGGAGAGCAGCAATCTATTCTGTGGCAGGCGG 1800

Db 1924 GCTTACAGCTGCGAGCAGCTGCCCACTTGGGAGAGCAGCAATCTATTCTGTGGCAGGCGG 1983

Qy 1801 CAGAGGCAAGTGAAGAGCTGAGCTGGCGGAGGAGCTGGCATGAGAGAGCCCTTTGAA 1860

Db 1984 CAGAGGCAAGTGAAGAGCTGAGCTGGCGGAGGAGCTGGCATGAGAGAGCCCTTTGAA 2043

Qy 1861 AAGCAGTTTAAAGCAGAGAGCTGCCAATGGAATTTGGAGAGAGCATCATGTGAGAGAC 1920

Db 2044 AAGCAGTTTAAAGCAGAGAGCTGCCAATGGAATTTGGAGAGAGCATCATGTGAGAGAC 2103

Qy 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTGTGCGGAGAGTGA 1980

Db 2104 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTGTGCGGAGAGTGA 2163

Qy 1981 ATCATGAGGTCTCC 1995

Db 2164 ATCATGAGGTCTCC 2178

RESULT 12

AB051487 4790 bp mRNA linear PRI 07-FEB-2001

LOCUS AB051487

DEFINITION Homo sapiens mRNA for KIAA1700 protein, partial cds.

ACCESSION AB051487.1 GI:12697944

VERSION AB051487.1

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases) Nagase, T., Kikuno, R., Hattori, A., Kondo, Y., Okumura, K. and Ohara, O. Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

TITLE Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro

JOURNAL DNA Res. 7 (6), 347-355 (2000)

MEDLINE 21082932

PUBMED 11214970

REFERENCE 2 (bases 1 to 4790)

AUTHORS Ohara, O., Nagase, T. and Kikuno, R.

TITLE Direct Submission

JOURNAL Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnaif@kazusa.or.jp, URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913, Fax: 81-438-52-3914)

FEATURES

source

1. .4790

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="Fj15353"

/note="vector: pBluescript SK plus"

1. .4790

/gene="KIAA1700"

<109. .2181

/gene="KIAA1700"

/note="Start codon is not identified."

/codon_start=1

/product="KIAA1700 protein"

/protein_id="BAB21791.1"

/db_xref="GI:12697945"

/translation="AEEHTTITSWQIKGGGRGLIVMAHEMIGTQIVTERVALL
ESGTEKVLIDSRPVEVNTSHILAININCKLMKRLQOQKVLITELIOHAKKV
DIDGOKVIVVDSODVASLSDFELTLGKLEKSPNSVHLIAGLPFSCFPL
CEGKSTVPTCISQPCLPVANTGPRILIPNLVLCORPDLANKMOONIGVLTARS
TCEKDFEIPESHFLRVNDSPCEKILPMLDVSVDIEKASNGECVYLHCLAGTARS
ATIAIAYIMKRMDSIDEAVRFVKEKRPISNPNFELGQLDEKIKNOTAGSGKS
KIKLILHEKPEPVPVASEGQKSETPSPCADSTSEAAQGPVHPASVSVSPVQ
PILBESPLVQALSGHLSADRLSDSNKLSRFSLDIKSVYSASVAAALHGSSES
ALEYKPSSTTLTGTKLQOPSPVOLESDPTSPSPKEKASI PKLQOTARPSPSKRS
LHSVTSSEGTQKRLSLSPHRSQSVEDVHNSPLPGLSTSOQHLTKSGCLKQHS
DLAPQSTSPSLTSSWTFATESHFTFSALITGGASVATSCSQLPTGDOVYSVR
RQKPDRASTRSRMSHESPFEKQFRKRSQOMEPGESIMSENRREBLRGVSQSFSFG
SMEILIEVS"

ORIGIN

Query Match 94.9%; Score 1893; DB 9; Length 4790;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATGGAATCAATTTGTTACTGAGAGGTGGTGGCTGTGAGAA 60

Db 184 ATGGCCCATGAGATGATGGAATCAATTTGTTACTGAGAGGTGGTGGCTGTGAGAA 243

Qy 61 AGTGAACGGAAAAAGTGTCTGCTAATTGATAGCCGGCCATTGTGGAATACATATCC 120

Db 244 AGTGAACGGAAAAAGTGTCTGCTAATTGATAGCCGGCCATTGTGGAATACATATCC 303

Qy 121 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 180

Db 304 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 363

Qy 181 GACAAAGTGTATTAATACAGAGCTCATCCAGCATTCAGCGAAACATAGGTTGATGAT 240

Db 364 GACAAAGTGTATTAATACAGAGCTCATCCAGCATTCAGCGAAACATAGGTTGATGAT 423

Qy 241 TGCAGTCAGAAAGTGTATTAATACAGATCAAGGCTCCCAAGATGTTCTCTCTTCA 300

Db 424 TGCAGTCAGAAAGTGTATTAATACAGATCAAGGCTCCCAAGATGTTCTCTCTTCA 483

Qy 301 GACTGTTTCTCACTGTAATCTTGGGTAATCTGAGAGAGCTTCAACTCTGTTCACTG 360

Db 484 GACTGTTTCTCACTGTAATCTTGGGTAATCTGAGAGAGCTTCAACTCTGTTCACTG 543

Qy 361 CTTGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420

Db 544 CTTGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 603

Qy 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTCTCTGTTGCCAATTTGGGCCAAC 480

Db 604 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTCTCTGTTGCCAATTTGGGCCAAC 663

Qy 481 CGAATTTCTCCCAATCTTATCTTGGCTCCAGCGAGATGCTCTCAACAGAGCTGATA 540

Db 664 CGAATTTCTCCCAATCTTATCTTGGCTCCAGCGAGATGCTCTCAACAGAGCTGATA 723

Qy 541 CAGCAAAATGGAGATGGTATGTTAATGCAAGCATACCTGTCAAAGCCTGACTT 600

Db 724 CAGCAAAATGGAGATGGTATGTTAATGCAAGCATACCTGTCAAAGCCTGACTT 783

Qy 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 660

Db 784 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAATTTTG 843

Qy 661 CCGTGTGAGCAATCACTAGATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTTT 720

Db 844 CCGTGTGAGCAATCACTAGATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTTT 903

Qy 721 CTAGTGCACCTGTTTACTGGGATCTCCGCTCCGCAACATGCTATCGCTACATCATG 780

Db 904 CTAGTGCACTGTTAGTGGGATCTCCGCTCCGCAACATCGTATGCTTACATCATG 963
Qy 781 AAGAGGATGGAACATGCTTTAGATGAAAGCTTACAGATTTTGGAAAAGAAAGCTTACT 840
Db 964 AAGAGGATGGAACATGCTTTAGATGAAAGCTTACAGATTTTGGAAAAGAAAGCTTACT 1023
Qy 841 ATATCTCCAAACCTTCAATTTTCTGGGCCAATCTCTGACTATGAGAAAGATTAAGAAC 900
Db 1024 ATATCTCCAAACCTTCAATTTTCTGGGCCAATCTCTGACTATGAGAAAGATTAAGAAC 1083
Qy 901 CAGACTGAGGATCAGGGCCAAAGAGCAATCAAGCTGTGCACTTGGAAAGCCAAAT 960
Db 1084 CAGACTGAGGATCAGGGCCAAAGAGCAATCAAGCTGTGCACTTGGAAAGCCAAAT 1143
Qy 961 GAACTGTGCTGCTGCTCAGAGGGTGGACAGAAAAGCGAGAGGCCCTCAGTCCACC 1020
Db 1144 GAACTGTGCTGCTGCTCAGAGGGTGGACAGAAAAGCGAGAGGCCCTCAGTCCACC 1203
Qy 1021 TGTGCTCCGACCTGTCTACTCAGAGGACAGAGCAAGAGGCCCTGTGCACTCCGCAAGCTG 1080
Db 1204 TGTGCTCCGACCTGTCTACTCAGAGGACAGAGCAAGAGGCCCTGTGCACTCCGCAAGCTG 1263
Qy 1081 CCCAGCGTGGCCGACGCTGACGCTGCTGTGAGAGCAGCCGCTGTGTAAGAGGCTTC 1140
Db 1264 CCCAGCGTGGCCGACGCTGACGCTGCTGTGAGAGCAGCCGCTGTGTAAGAGGCTTC 1323
Qy 1141 AGTGGGCTGACCTGTCCGAGACAGAGCTGAAAGACAGCAATAGCTCAAGGCTTCTTC 1200
Db 1324 AGTGGGCTGACCTGTCCGAGACAGAGCTGAAAGACAGCAATAGCTCAAGGCTTCTTC 1383
Qy 1201 TCTCTGATATCAAAATCAGTTTCAATATTCAGCCAGACATGAGCATCTTACATGCTTC 1260
Db 1384 TCTCTGATATCAAAATCAGTTTCAATATTCAGCCAGACATGAGCATCTTACATGCTTC 1443
Qy 1261 TCTCTCATCAGAAAGTGTGGAATCTCAAACTTCCACTCTCTGATGAGACCAAC 1320
Db 1444 TCTCTCATCAGAAAGTGTGGAATCTCAAACTTCCACTCTCTGATGAGACCAAC 1503
Qy 1321 AAGCTATGCGAGTTCTCCCTGTTTCAGAACTATCCGAGACAGCTCCGAAACAGTCTCT 1380
Db 1504 AAGCTATGCGAGTTCTCCCTGTTTCAGAACTATCCGAGACAGCTCCGAAACAGTCTCT 1563
Qy 1381 GATTAAGAGAGAACCCAGCATCTCCCAAGAAAGCTGACAGCCGCAAGGCTTTCAGACAGCCAG 1440
Db 1564 GATTAAGAGAGAACCCAGCATCTCCCAAGAAAGCTGACAGCCGCAAGGCTTTCAGACAGCCAG 1623
Qy 1441 AGCAAGCGATTGATTCGGTCAGAAACAGAGAGTGGACAGCCGCAAGGCTTTCAGTTC 1500
Db 1624 AGCAAGCGATTGATTCGGTCAGAAACAGAGAGTGGACAGCCGCAAGGCTTTCAGTTC 1683
Qy 1501 TCTCCACTGATCGAAGTGGAGCGTGAAGCAATTAACACAGAGCTTCTTTTCGAC 1560
Db 1684 TCTCCACTGATCGAAGTGGAGCGTGAAGCAATTAACACAGAGCTTCTTTTCGAC 1743
Qy 1561 CTTTCCACGAGCAGAGACACTTCAAGAAAGTCTGCTGCTGAGCTTAAAGGCTGAGAC 1620
Db 1744 CTTTCCACGAGCAGAGACACTTCAAGAAAGTCTGCTGCTGAGCTTAAAGGCTGAGAC 1803
Qy 1621 TGGGATATCTTGGCCCCCAGACCTTACCCCTTCCGAGACAGAGAGCTGGATTTTGGC 1680
Db 1804 TGGGATATCTTGGCCCCCAGACCTTACCCCTTCCGAGACAGAGAGCTGGATTTTGGC 1863
Qy 1681 AAGAGATCTTCACTTCTACTCTGCTGACGCAATCTAAGAGGAGCTGCAATCTACTCT 1740
Db 1864 AAGAGATCTTCACTTCTACTCTGCTGACGCAATCTAAGAGGAGCTGCAATCTACTCT 1923
Qy 1741 GCTTACAGCTGACGAGCTGCGCACTTGGCGAGACCAAGTCTATCTGTGCGAGAGGCG 1800
Db 1924 GCTTACAGCTGACGAGCTGCGCACTTGGCGAGACCAAGTCTATCTGTGCGAGAGGCG 1983
Qy 1801 CAGAAAGCAAGTGAAGAGCTGACTCGCGGAGAGCTGGCATGAAGAGAGCCCTTTGAA 1860

Db 1984 CAGAAAGCAAGTGAAGAGCTGACTGCTCGCGCGAGAGCTGGCATGAAGAGAGCCCTTTGAA 2043
Qy 1861 AAGCACTTTTAAACGAGAAAGCTGCGCAAAATGAAATTTGAGAGAGCATGTCAGAGAAC 1920
Db 2044 AAGCACTTTTAAACGAGAAAGCTGCGCAAAATGAAATTTGAGAGAGCATGTCAGAGAAC 2103
Qy 1921 AGGTACCGGAAAGAGCTGGGAAAGTGGGCACTGCTAGCTTTTTCGGGACGATGAA 1980
Db 2104 AGGTACCGGAAAGAGCTGGGAAAGTGGGCACTGCTAGCTTTTTCGGGACGATGAA 2163
Qy 1981 ATCATTGAGGTCTCC 1995
Db 2164 ATCATTGAGGTCTCC 2178
RESULT 13
AX180875
LOCUS AX180875 2732 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 2 from Patent WO0146394.
ACCESSION AX180875
VERSION AX180875.1 GI:15132703
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Plowman, G.D., Martinez, R., Whyte, D., Manning, G., Sudarsanam, S.,
Hill, R.J. and Planagan, P.
TITLE Mammalian protein phosphatases
JOURNAL Patent: WO 0146394-A 2 28-JUN-2001;
Sugen, Inc. (US)
FEATURES
source location/Qualifiers
1..2732
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 92.3%; Score 1842; DB 6; Length 2732;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGGCCCATGAGATGATTTGGAACATAATTGTTACTGAGAGGTGGGCTGCTGGAA 60
Db 538 ATGGCCCATGAGATGATTTGGAACATAATTGTTACTGAGAGGTGGGCTGCTGGAA 597
Qy 61 AGTGAACGGAAGAGTGTCTGCTTAATTGATAGCCGCGCAATTTGGAATCAATACATCC 120
Db 598 AGTGAACGGAAGAGTGTCTGCTTAATTGATAGCCGCGCAATTTGGAATCAATACATCC 657
Qy 121 CACATTTTGAAGCCATTAATATCACTGCTCAAGCTTATGAAGGAAGTTGCAACAG 180
Db 658 CACATTTTGAAGCCATTAATATCACTGCTCAAGCTTATGAAGGAAGTTGCAACAG 717
Qy 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTGACATTCAT 240
Db 718 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTGACATTCAT 777
Qy 241 TGCAGTCAGAAAGTGTGATTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
Db 778 TGCAGTCAGAAAGTGTGATTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 837
Qy 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGAGCTTCAACTCTGTTACCTG 360
Db 838 GACTGTTTCTCACTGATCTTCTGGGTAACTGAGAAAGAGCTTCAACTCTGTTACCTG 897
Qy 361 CTTCAGAGTGGTGGTGGTCTGAGTCTCTGTTTTCCTGAGCTGTTGAAGAAATTC 420
Db 898 CTTCAGAGTGGTGGTGGTCTGAGTCTCTGTTTTCCTGAGCTGTTGAAGAAATTC 957
Qy 421 ACTCTAGTCCCTTACCTGCAATTTCTGAGCTTGTCTTACCTGTTGCAACATTTGGGCAAC 480

Db 958 ACTCTAGTCCCTACCTGATCTTCTCAGACCTTGCTTACCTGTGCAAACTATGGGCAACC 1017
Qy 481 CGAATCTTCCCAATCTTATCTTGGCTGCGACGAGATGTCTCAACAAAGAGCTGATA 540
Db 1018 CGAATCTTCCCAATCTTATCTTGGCTGCGACGAGATGTCTCAACAAAGAGCTGATA 1077
Qy 541 CAGCAGATGGATGGATGTTATGTTAAATGAGCAGTATCTGTCCAAAGCTGACTT 600
Db 1078 CAGCAGATGGATGGATGTTATGTTAAATGAGCAGTATCTGTCCAAAGCTGACTT 1137
Qy 601 ATCCCGAGTCTGATTTCTGCGGTGCTGTGATGACAGCTTTTGTGAGAAATTTTG 660
Db 1138 ATCCCGAGTCTGATTTCTGCGGTGCTGTGATGACAGCTTTTGTGAGAAATTTTG 1197
Qy 661 CCGTGGTGGACAAATCAGTATGATTTTATTTGAGAAAGCAAAAGCTTCCATGATGTT 720
Db 1198 CCGTGGTGGACAAATCAGTATGATTTTATTTGAGAAAGCAAAAGCTTCCATGATGTT 1257
Qy 721 CTAGTGAATGTTAGTGGGATCTCCGCTCCGCGACATGGCTATGCTGATCATG 780
Db 1258 CTAGTGAATGTTAGTGGGATCTCCGCTCCGCGACATGGCTATGCTGATCATG 1317
Qy 781 AAGAGATGACATGTTCTTATGATGAGCTTACAGATTTGTGAGAAAGAAAGACCTACT 840
Db 1318 AAGAGATGACATGTTCTTATGATGAGCTTACAGATTTGTGAGAAAGAAAGACCTACT 1377
Qy 841 ATATCTCCAAATCTCAATTTTCTGGGCAATCTCTGAGCTATGAGAAAGATTAAGAC 900
Db 1378 ATATCTCCAAATCTCAATTTTCTGGGCAATCTCTGAGCTATGAGAAAGATTAAGAC 1437
Qy 901 CAGACTGAGACATCAGGCGCAAGAGCAAACTCAGCTGTGACCTGGAGAGCCAAAT 960
Db 1438 CAGACTGAGACATCAGGCGCAAGAGCAAACTCAGCTGTGACCTGGAGAGCCAAAT 1497
Qy 961 GAACCTGTCCCTGCTGCTCAGAGGCTGACAGAAAAAGCAGCCCTCAGTCCACC 1020
Db 1498 GAACCTGTCCCTGCTGCTCAGAGGCTGACAGAAAAAGCAGCCCTCAGTCCACC 1557
Qy 1021 TGTGCGAATCTGCTCAGAGGCTGACAGAAAAAGCAGCCCTCAGTCCACC 1080
Db 1558 TGTGCGAATCTGCTCAGAGGCTGACAGAAAAAGCAGCCCTCAGTCCACC 1617
Qy 1081 CCGAGGCTGCGACGCGTGCCTGCTTGAAGAGACCCCGCTGATACAGGGCTC 1140
Db 1618 CCGAGGCTGCGACGCGTGCCTGCTTGAAGAGACCCCGCTGATACAGGGCTC 1677
Qy 1141 AGTGGCTGCACTGTCTCCGACAGAGCTGGAAGACGAATAAGCTCAAGCGTCTTC 1200
Db 1678 AGTGGCTGCACTGTCTCCGACAGAGCTGGAAGACGAATAAGCTCAAGCGTCTTC 1737
Qy 1201 TCTCTGATATCAATCAGTTTCAATATTCAGCCAGCAGTGCATCTTCAATGCTTC 1260
Db 1738 TCTCTGATATCAATCAGTTTCAATATTCAGCCAGCAGTGCATCTTCAATGCTTC 1797
Qy 1261 TCTCTCATCAGAAATGTTTGAATATCTCAAACTTCTGATGAGGAGCAAC 1320
Db 1798 TCTCTCATCAGAAATGTTTGAATATCTCAAACTTCTGATGAGGAGCAAC 1857
Qy 1321 AAGCTATGCAAGTTTCTCCCTGTTCAAGAACTATCGAGAGAGCTCCCGAAACAGTCT 1380
Db 1858 AAGCTATGCAAGTTTCTCCCTGTTCAAGAACTATCGAGAGAGCTCCCGAAACAGTCT 1917
Qy 1381 GATTAAGAGAGAGCAGCATCCCAAGAGCTGACAGCCGAGCGCTTCAAGCAGCAG 1440
Db 1918 GATTAAGAGAGAGCAGCATCCCAAGAGCTGACAGCCGAGCGCTTCAAGCAGCAG 1977
Qy 1441 AGCAAGGATTTGATTTGGTTCAGAAACAGCAGAGTGCACCGCCAGAGAGTCTCTTTA 1500
Db 1978 AGCAAGGATTTGATTTGGTTCAGAAACAGCAGAGTGCACCGCCAGAGAGTCTCTTTA 2037
Qy 1501 TCTCTCACTGATGAGAGTGGAGAGTGGAGAGCAATTAACAACAGCTTCTTTTCGG 1560
Db 2038 TCTCTCACTGATGAGAGTGGAGAGTGGAGAGCAATTAACAACAGCTTCTTTTCGG 2097

Qy 1561 CTTTCCACGAGCAGAGCAGCCTCAGAGAGTCTGCTGAGGCTTAAGGGCTGGCAC 1620
Db 2098 CTTTCCACGAGCAGAGCAGCCTCAGAGAGTCTGCTGAGGCTTAAGGGCTGGCAC 2157
Qy 1621 TCGGATATCTTGGCCCCCAGACCTTACCTTCCCTGACGACGAGCTGATTTTGGC 1680
Db 2158 TCGGATATCTTGGCCCCCAGACCTTACCTTCCCTGACGACGAGCTGATTTTGGC 2217
Qy 1681 ACGAGTCTCAACTTCTTCTGCTGCTGACCATCTACGAGGAGCTGCCAATCTCT 1740
Db 2218 ACGAGTCTCAACTTCTTCTGCTGCTGACCATCTACGAGGAGCTGCCAATCTCT 2277
Qy 1741 GCTTACAGTGAAGCAGAGTGCCTTCCGAGAGCAAGTCTATCTGCGCAGGCG 1800
Db 2278 GCTTACAGTGAAGCAGAGTGCCTTCCGAGAGCAAGTCTATCTGCGCAGGCG 2337
Qy 1801 CAGAGCCAAAGTGAAGAGCTGCTGCGCGAGAGCTGACATGAAGAGAGCCCTTTGA 1860
Db 2338 CAGAGCCAAAGTGAAGAGCTGCTGCGCGAGAGCTGACATGAAGAGAGCCCTTTGA 2397
Qy 1861 AAGCAGTTAAACGAGAGCTGCCAATGGAATTTGAGAGAGCATCTGTCAGAGAC 1920
Db 2398 AAGCAGTTAAACGAGAGCTGCCAATGGAATTTGAGAGAGCATCTGTCAGAGAC 2457
Qy 1921 AGGTACCGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTCTTGGGCGAGCAGAA 1980
Db 2458 AGGTACCGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTCTTGGGCGAGCAGAA 2517
Qy 1981 ATCATTGAGTCTCC 1995
Db 2518 ATCATTGAGTCTCC 2532

RESULT 14
AX405700
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

1. 3104
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
159. .2156
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD34807.1"
/db_xref="GI:21438840"
/db_xref="EMBL:CAD34807"
/translation="MAHEMIGQIVTERVALLLEGTEKLVLLIDSRPFVEYNTSHILE
AININSLKMRRLQDDKVLITELIOHSAKHVYDIDCSQKVVVYDSSQDVASLSBDC
FLTVLLGKLEKSPNVHLAAGFAFESRCFPGICGSKFLVPTCISOPCLPVAATIGPT
RLPVLNLYGECORDIVNKEIMQONGYVYNAENTCPKPFIPESHFLRPVNDSECEK
ILPMLDKSDTFLBKAKSNGCTVYHCLAGISASATLAIYIMKRDMSLDEAYRPVKE
KRPITSPPNENFLGQLDYEKTIKQNTGASGPYSKULHLERPNBEPVAVSFGQKSE
TLPSPCADSATSEBAQRPVAPVAPVSPQPSLLEDSPLVQALSGHLASADRLED
SNKLRKSFSLDIKSYSAASMAASLHGFSSESDALEYKPTLLDGTNLCOFSYVQES
LSBQRPETSPDEBASIPKLIQTAPSPDSQKRLHSVRSSGTQKORSLSLPSLRGGS
VBDNYHTSFLRSLTSSOOLHTSAGLIGKGMHSDILAPOTSRPSLTSMSVPTESHF
YASALYGGASAYSISCSQLPTCGDQVYSVRROKPSPRADSRSRMSHSESPKQPK

ORIGIN

RSCQMFGEISIMSENRSREBELKVGSGSSFSGSMELIEVS"

Query Match 92.3%; Score 1842; DB 6; Length 3104;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 ATGSGCCATGAGATGATGTAAGTCAATTTGTAAGAGAGTGTGGCTGCTGAGAA 60
DB 159 ATGSGCCATGAGATGATGTAAGTCAATTTGTAAGAGAGTGTGGCTGCTGAGAA 218
QY 61 AGTGAACGGAAAAAGTGTGTAATGTAAGCCGGCCATTGTGGAATACATATAC 120
DB 219 AGTGAACGGAAAAAGTGTGTAATGTAAGCCGGCCATTGTGGAATACATATAC 278
QY 121 CACATTTTGAAGCCATTAATCAACTGTCCAGCTTAAGAGCGAAGTTGCAACG 180
DB 279 CACATTTTGAAGCCATTAATCAACTGTCCAGCTTAAGAGCGAAGTTGCAACG 338
QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGAT 240
DB 339 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGAT 398
QY 241 TGCAGTCAGAAAGTTGATTAACGATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 300
DB 399 TGCAGTCAGAAAGTTGATTAACGATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 458
QY 301 GACTGTTTTCTCACTGTACTCTTGCGGTAACTGGAAGAAGCTTCAACTCTGTTCA 360
DB 459 GACTGTTTTCTCACTGTACTCTTGCGGTAACTGGAAGAAGCTTCAACTCTGTTCA 518
QY 361 CTTCGAGGTGGGTTTGTGAGTCTCTGTTGTTTCCCGGCTCTGTGAAGGAATCC 420
DB 519 CTTCGAGGTGGGTTTGTGAGTCTCTGTTGTTTCCCGGCTCTGTGAAGGAATCC 578
QY 421 ACTCTAGTCCCTAAGCTGATTTCTCAGCCTTGCTTAAGTGGCCAACTTGGCCAA 480
DB 579 ACTCTAGTCCCTAAGCTGATTTCTCAGCCTTGCTTAAGTGGCCAACTTGGCCAA 638
QY 481 CGAATTTCTCCCAATCTTTATCTTGCGTGCACGAGATGTCCTCAACAAGAGCTGANA 540
DB 639 CGAATTTCTCCCAATCTTTATCTTGCGTGCACGAGATGTCCTCAACAAGAGCTGANA 698
QY 541 CAGCAGATGGAGTTGATTAATGTTAAATGCGAGTATACCTGTCCAAAGCCTGACTT 600
DB 699 CAGCAGATGGAGTTGATTAATGTTAAATGCGAGTATACCTGTCCAAAGCCTGACTT 758
QY 601 ATCCCGAGTCTGATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG 660
DB 759 ATCCCGAGTCTGATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG 818
QY 661 CCGTGTGTGACAAATCAGTGAATTTGATTTGAGAAAAGCAAAAGCCTCCATGATGTT 720
DB 819 CCGTGTGTGACAAATCAGTGAATTTGATTTGAGAAAAGCAAAAGCCTCCATGATGTT 878
QY 721 CTAGTCACTGTTTACGTGGATCTCCGCTCCGACCATGCTATGCTATCATCATG 780
DB 879 CTAGTCACTGTTTACGTGGATCTCCGCTCCGACCATGCTATGCTATCATCATG 938
QY 781 AAGAGATGAGCAATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 840
DB 939 AAGAGATGAGCAATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACT 998
QY 841 ATATCTCCAAATCTCAATTTTCTGGGCAACTCTGACTATGAGAAAGATTAAGAAC 900
DB 999 ATATCTCCAAATCTCAATTTTCTGGGCAACTCTGACTATGAGAAAGATTAAGAAC 1058
QY 901 CAGACTGAGCATCAGGGCCAAAGACAACTCAAGCTGTGACCTGAGAGAGCAAAAT 960
DB 1059 CAGACTGAGCATCAGGGCCAAAGACAACTCAAGCTGTGACCTGAGAGAGCAAAAT 1118
QY 961 GAACTGTCTCTGCTGTCTCAGAGGTGACAGAAAAAGGAGACGCCCTCAGTCCACCC 1020
```

```
DB 1119 GAACTGTCTCTGCTGTCTCAGAGGTGACAGAAAAAGGAGACGCCCTCAGTCCACCC 1178
QY 1021 TGTGCCGACTGTGTACTCTCAGAGGACAGACAAAGCCCTGTGATTCGCCGACGGT 1080
DB 1179 TGTGCCGACTGTGTACTCTCAGAGGACAGACAAAGCCCTGTGATTCGCCGACGGT 1238
QY 1081 CCCAGCGGCCAGAGGTGACGGCTGTGCTGTGAGAGACAGCCCGCTGATCAGGCGCTC 1140
DB 1239 CCCAGCGGCCAGAGGTGACGGCTGTGCTGTGAGAGACAGCCCGCTGATCAGGCGCTC 1298
QY 1141 AGTGGGCTGACCTGTCCGACAGAGCTGTGAGAGACAGCAATTAAGCTCAAGCTTCTTC 1200
DB 1239 AGTGGGCTGACCTGTCCGACAGAGCTGTGAGAGACAGCAATTAAGCTCAAGCTTCTTC 1358
QY 1201 TCTCTGATATCAATTCAGTTTCAATTCAGCAGACATGAGCATCTCTTACATGGCTTC 1260
DB 1359 TCTCTGATATCAATTCAGTTTCAATTCAGCAGACATGAGCATCTCTTACATGGCTTC 1418
QY 1261 TCTCTGATGAAAGATGCTTTGGAATTAACAACCTTCCACTACTCTGGATGGAACAAC 1320
DB 1419 TCTCTGATGAAAGATGCTTTGGAATTAACAACCTTCCACTACTCTGGATGGAACAAC 1478
QY 1321 AAGCTATGCAAGTTCTCCCTGTTTCAAGAACTATTCGAGAGACATCCCGAAAACAGTCT 1380
DB 1479 AAGCTATGCAAGTTCTCCCTGTTTCAAGAACTATTCGAGAGACATCCCGAAAACAGTCT 1538
QY 1381 GATTAAGAGAGAACAGAGATCCCGAAGAGCTGACACCGCCAGGCTTTCAGACAGCAG 1440
DB 1539 GATTAAGAGAGAACAGAGATCCCGAAGAGCTGACACCGCCAGGCTTTCAGACAGCAG 1598
QY 1441 AGCAAGCATGTCATTCGGTTCAGAACCCAGCAGAGTGGACCCGACAGAGTCCCTTTTA 1500
DB 1599 AGCAAGCATGTCATTCGGTTCAGAACCCAGCAGAGTGGACCCGACAGAGTCCCTTTTA 1658
QY 1501 TCTCCATGTCATCGAAGTGGAGGCTGTGAGACAAATTAACACACAGCTTCTTTTCCGC 1560
DB 1659 TCTCCATGTCATCGAAGTGGAGGCTGTGAGACAAATTAACACACAGCTTCTTTTCCGC 1718
QY 1561 CTTTCACACAGCAGACAGACCTCAAGAGTGTGCTGCGCTTAAGGGCTGGAC 1620
DB 1719 CTTTCACACAGCAGACAGACCTCAAGAGTGTGCTGCGCTTAAGGGCTGGAC 1778
QY 1621 TCGATATCTTGGCCCCCAGACCTTAACCCCTTCCCTGACAGCAGAGCTGTATTTTGGC 1680
DB 1779 TCGATATCTTGGCCCCCAGACCTTAACCCCTTCCCTGACAGCAGAGCTGTATTTTGGC 1838
QY 1681 ACAGAGTCTCAACTTCTACTGTGCTGACGATCTACGAGGAGGAGGAGGAGGAGTACTCT 1740
DB 1839 ACAGAGTCTCAACTTCTACTGTGCTGACGATCTACGAGGAGGAGGAGGAGGAGTACTCT 1898
QY 1741 GCTTACAGTGAAGCAGCTGCCCACTTGGCGGAGCAAGTCTATTTGTGCGCAGGCGG 1800
DB 1899 GCTTACAGTGAAGCAGCTGCCCACTTGGCGGAGCAAGTCTATTTGTGCGCAGGCGG 1958
QY 1801 CAGAGCCAAAGTGAAGAGCTGACTCGCGCGGAGAGCTGAGAGAGAGAGCCCTTTTGA 1860
DB 1959 CAGAGCCAAAGTGAAGAGCTGACTCGCGCGGAGAGCTGAGAGAGAGAGAGCCCTTTTGA 2018
QY 1861 AAGCAGTTTAAACGAGAAAGCTGCCAAATGGAATTTGAGAGAGATCATGTCAAGAAC 1920
DB 2019 AAGCAGTTTAAACGAGAAAGCTGCCAAATGGAATTTGAGAGAGATCATGTCAAGAAC 2078
QY 1921 AGGTCAAGGGAAGAGCTGGGGAAGTGGGAGTCAAGTCAAGCTTTTGGGCGAGATGAA 1980
DB 2079 AGGTCAAGGGAAGAGCTGGGGAAGTGGGAGTCAAGTCAAGCTTTTGGGCGAGATGAA 2138
QY 1981 ATCATTGAGTCTCC 1995
DB 2139 ATCATTGAGTCTCC 2153
```

RESULT 15
AX713989

Db 1796 GCTTACAGCTGCAGCCAGCTGCCACCTTGGGAGAGACCAATGCTATTCTGTGGCAGCGG 1855

QY 1801 CAGAACCCAGATGACAGAGTGACTGCGCGCGGAGCTGGCATGAAGACCCCTTTGAA 1860

Db 1856 CAGAACCCAGATGACAGAGTGACTGCGCGCGGAGCTGGCATGAAGACCCCTTTGAA 1915

QY 1861 AAGCATTTAAACGACAGAGCTGCCAAATGGAATTTGAGAGAGCATGTCAGAGAC 1920

Db 1916 AAGCATTTAAACGACAGAGCTGCCAAATGGAATTTGAGAGAGCATGTCAGAGAC 1975

QY 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTACCTTTTGGGAGCATGAA 1980

Db 1976 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTTACCTTTTGGGAGCATGAA 2035

QY 1981 ATCATTTAGTCTCC 1995

Db 2036 ATCATTTAGTCTCC 2050

RESULT 16
AK055973

LOCUS 2102 bp mRNA linear PRI 01-AUG-2002

DEFINITION Homo sapiens CDNA FLJ31411 f1s, clone NT2NE2000214, moderately similar to DUAL SPECIFICITY PROTEIN PHOSPHATASE 8 (EC 3.1.3.48).

ACCESSION AK055973

VERSION AK055973.1 GI:16550835

KEYWORDS oligo capping; f1s (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukushima, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Ito, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Watsushima, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahata, K., Masuno, Y., Nagai, K. and Isogai, T.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2102)

AUTHORS Isogai, T., Otsuki, T. and Sugiyama, T.

TITLE Direct Submission

JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genom@csb.hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES

source

1.2102

location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT2NE2000214"

/cell_line="NT2"

/cell_type="teratocarcinoma"

/clone_lib="NT2NE2"

/note="Cloning vector: pME18FLJ3-mRNA from NT2 neuron after the differentiation of NT2 neuronal precursor cells."

56..2053

/note="unnamed protein product"

/codon_start=1

/protein_id="BAB71060.1"

CDS

/db_xref="GI:16550835"

/translation="MAHEMIGQIYTERLVALLBSETEKVLIDSRPEVNTSHILE
AIVNIGSKMRRLQDDVLTTELIOHSAKHREVDIGQKVVYDIOSSDVALSDC
FLTVLGLKERSFNSVHLLAGFAEFSRPELCEGKSLVPTCTIOSPLVPAVNGPT
LILPNLYGCDRDVNLKELMOQNGIYVLNANTRCPKDPFIPESHFLVPAVDSCEK
ILPMDKXVDVELEKAKNSGCVLNGIERSATIAIAYIKRMKMDLDEAYRFEKE
KRPTSPNPNFLGOLLIDYEKKIKNOTGASGPKSRKLLHLEKPNRPVAVSGGKSE
TPLSPCADSATSSEAGORPVHPASVSPVSPVLPDPLVQALSGLHLISADLEL
SNKLRSLDIKSVSYSAASAHLHGSSSEDALEYRPETTLDTGNTKLCPSVOE
LSEOTPEPSPKREASIPKIQOTAPSDSQRHLHSVRSISSGTRQSLSPVLRSGS
VEDNYHTSFLFGLSTSOQLTKSAGLGLKWHMSIDLAPQTSPLTSSWYPTESHF
YSASAIYGSASYSVASCOLPTCGDQVYSVRKRPEDRADRSRMWESFEKQFK
RRSCMERGESIMSENRERELGKVSQSSPSGSEILIEVS"

Query Match 89.8%; Score 1791; DB 9; Length 2102;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATTTGAACTCAATTTGTAAGAGAGTGGTCTGCTGAGAA 60

Db 56 ATGGCCCATGAGATGATTTGAACTCAATTTGTAAGAGAGTGGTCTGCTGAGAA 115

QY 61 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCCGGCATTGTTGAAATACAAATACATCC 120

Db 116 AGTGAACGGAAGAAAGTCTGCTAATTGATAGCCGGCATTGTTGAAATACAAATACATCC 175

QY 121 CACATTTTGAAGACCATTAATATCAAGCTCTCCAAAGCTTTATGAAAGCGAGTTGCAACG 180

Db 176 CACATTTTGAAGACCATTAATATCAAGCTCTCCAAAGCTTTATGAAAGCGAGTTGCAACG 235

QY 181 GACAAAGTGTATATACAGAGCTCATCCAGATTGAGGAAACATPAGTTGACATTGAT 240

Db 236 GACAAAGTGTATATACAGAGCTCATCCAGATTGAGGAAACATPAGTTGACATTGAT 295

QY 241 TGCAGTCAGAAAGTTGATGATTAAGATCAAGAGTCCCAAGATGTTGCTCTCTCTTCA 300

Db 296 TGCAGTCAGAAAGTTGATGATTAAGATCAAGAGTCCCAAGATGTTGCTCTCTCTTCA 355

QY 301 GACTGTTTTCACACTGTACTCTGCTGATTAACATCGAAGAGACCTTCACTCTGTTCACTG 360

Db 356 GACTGTTTTCACACTGTACTCTGCTGATTAACATCGAAGAGACCTTCACTCTGTTCACTG 415

QY 361 CTTCGAGGTGGATTTGCTGAGATTCCTGCTGTTTCCCTGCTCTGTGAGAGAAATCC 420

Db 416 CTTCGAGGTGGATTTGCTGAGATTCCTGCTGTTTCCCTGCTCTGTGAGAGAAATCC 475

QY 421 ACTTAGTCCCTACCTGATTTCTCAGCCTTCTTACCTGTTGGCAACATTTGGGCAACC 480

Db 476 ACTTAGTCCCTACCTGATTTCTCAGCCTTCTTACCTGTTGGCAACATTTGGGCAACC 535

QY 481 CGAATTCCTCCCATCTTATCTTGGCTGCTGAGAGATGTCCTCAACAGAGCTGATA 540

Db 536 CTAATTCCTCCCATCTTATCTTGGCTGCTGAGAGATGTCCTCAACAGAGCTGATA 595

QY 541 CAGCAGAAATGGAGATTTGATGTTAATGTCAGACTATACCTGCAAAAGCTGACCTT 600

Db 596 CAGCAGAAATGGAGATTTGATGTTAATGTCAGACTATACCTGCAAAAGCTGACCTT 655

QY 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660

Db 656 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 715

QY 661 CCGGTGTTGGAACAATAGTATGATTTCAATGAGAAAGCAAAAGCTCCATGATGTGT 720

Db 716 CCGGTGTTGGAACAATAGTATGATTTCAATGAGAAAGCAAAAGCTCCATGATGTGT 775

QY 721 CTATGACATGTTTATGCTGGGATCTCCCGCTCCGCAACATCGTATGCGCTATCATATG 780

Db 776 CTATGACATGTTTATGCTGGGATCTCCCGCTCCGCAACATCGTATGCGCTATCATATG 835

QY 781 AAGAGATGACATGTTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTACT 840

Db 836 AAGAGATGAGATGCTTTAGATGAGCTTACAGATTTGTGAAAGAAAAAGACTACT 895
QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGGACTATGAGAAAGATTAAAGAC 900
Db 896 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGGACTATGAGAAAGATTAAAGAC 955
QY 901 CAGACTGAGACATCAGGGCCAAAGAGCAATCAAGCTGTGCACTGTGAAAGCAAAAT 960
Db 956 CAGACTGAGACATCAGGGCCAAAGAGCAATCAAGCTGTGCACTGTGAAAGCAAAAT 1015
QY 961 GAACCTGTCTCTGTCTCTCAAGGGTGAACAGAAAAAGCAAGCCCTCACTCCACC 1020
Db 1016 GAACCTGTCTCTGTCTCTCAAGGGTGAACAGAAAAAGCAAGCCCTCACTCCACC 1075
QY 1021 TGAGCGCACTCTGTCTCTCAAGGGGAGAGCAAAAGCCGTGCACTCCGCAAGGTG 1080
Db 1076 TGAGCGCACTCTGTCTCTCAAGGGGAGAGCAAAAGCCGTGCACTCCGCAAGGTG 1135
QY 1081 CCCAGCGTCCAGCGCTGACAGCCGTGCTGTAGAGAGACAGCCGCTGTACAGAGGCTC 1140
Db 1136 CCCAGCGTCCAGCGCTGACAGCCGTGCTGTAGAGAGACAGCCGCTGTACAGAGGCTC 1195
QY 1141 AGTGGGCTGCACTGTCTCCAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
Db 1196 AGTGGGCTGCACTGTCTCCAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1255
QY 1201 TCTCTGGATATCAATTCAGTTTATTCAGCCAGAGATGGACATCTTCAATGGCTTC 1260
Db 1256 TCTCTGGATATCAATTCAGTTTATTCAGCCAGAGATGGACATCTTCAATGGCTTC 1315
QY 1261 TCCTCATCAGAAAGATGCTTTGGAATACTAACAACCTTCACTCTGTAGTGGAGCAAC 1320
Db 1316 TCCTCATCAGAAAGATGCTTTGGAATACTAACAACCTTCACTCTGTAGTGGAGCAAC 1375
QY 1321 AAGCTATGCCAGTTTCTCCCTGTTCAAGAACTATCGAGAGACATCTCCGAAACAGTCT 1380
Db 1376 AAGCTATGCCAGTTTCTCCCTGTTCAAGAACTATCGAGAGACATCTCCGAAACAGTCT 1435
QY 1381 GATTAAGAGAGAACCGACATCCCGAAGACCTGACAGCCGCTTACAGACCCAG 1440
Db 1436 GATTAAGAGAGAACCGACATCCCGAAGACCTGACAGCCGCTTACAGACCCAG 1495
QY 1441 AGCAAGGATTTGATTTGCTGTCAGAAACAGCAGAGTGGACCCGCAAGAGTCCCTTTTA 1500
Db 1496 AGCAAGGATTTGATTTGCTGTCAGAAACAGCAGAGTGGACCCGCAAGAGTCCCTTTTA 1555
QY 1501 TCTCACTGATGAAAGTGGAGCGTGGAGACATTAACAACAGCTTCTTTTCGAG 1560
Db 1556 TCTCACTGATGAAAGTGGAGCGTGGAGACATTAACAACAGCTTCTTTTCGAG 1615
QY 1561 CTTTCAACGAGCAGACGACCTCAAGAACTCTGCTGGCTTGAAGGGCTGGCAC 1620
Db 1616 CTTTCAACGAGCAGACGACCTCAAGAACTCTGCTGGCTTGAAGGGCTGGCAC 1675
QY 1621 TCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGAGCTGTTATTTTGGC 1680
Db 1676 TCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGAGCTGTTATTTTGGC 1735
QY 1681 ACAGAGTCTCTACAGATTTCTACTGTGCTCAAGCATTAACGAGAGCAAGTGGCACTACT 1740
Db 1736 ACAGAGTCTCTACAGATTTCTACTGTGCTCAAGCATTAACGAGAGCAAGTGGCACTACT 1795
QY 1741 GCTTACAGCTGCAAGCAGCTGCCCACTTGGGAGAGCAAGTCTTATTTCTGTGGCAGAGCG 1800
Db 1796 GCTTACAGCTGCAAGCAGCTGCCCACTTGGGAGAGCAAGTCTTATTTCTGTGGCAGAGCG 1855
QY 1801 CAGAAAGCAAGTACAGAGCTGAATCGCGCGAGAGCTGAGTAAAGAGAGCCCTTTGAA 1860
Db 1856 CAGAAAGCAAGTACAGAGCTGAATCGCGCGAGAGCTGAGTAAAGAGAGCCCTTTGAA 1915
QY 1861 AAGCAATTTAAACGAGAGCTGCCAAATGAAATTTGAGAGAGCATCATGTCAAGAAAC 1920
Db 1916 AAGCAATTTAAACGAGAGCTGCCAAATGAAATTTGAGAGAGCATCATGTCAAGAAAC 1975

QY 1921 AGGTACCGGGAAGAGAGCTGGGGAAGTGGGCACTAGCTTATGCTTTTGGGCAAGATGAA 1980
Db 1976 AGGTACCGGGAAGAGAGCTGGGGAAGTGGGCACTAGCTTATGCTTTTGGGCAAGATGAA 2035
QY 1981 ATCATTGAGTCTCC 1995
Db 2036 ATCATTGAGTCTCC 2050

RESULT 17
AX482372 5111 bp DNA linear PAT 16-AUG-2002
LOCUS AX482372
DEFINITION Sequence 41 from Patent WO02057460.
ACCESSION AX482372
VERSION AX482372.1 GI:22316945
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Todderud, C.G., Bol, D., Finger, J., Lee, L., Nelson, T., Schieven, G.,
Suchard, S., Banas, D., Bassolino, D., Feder, J., Kyatek, S.,
Mcatee, P., Mintier, G., Siemers, N., Jackson, D.G. and Ramanathan, C.
Polynucleotides encoding human phosphatases
Patent: WO 02057460-A 41 25-JUL-2002;
SQUIBB BRISTOL MYERS CO (US)
location/Qualifiers
1..5111
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES
source

ORIGIN
Query Match 89.8%; Score 1791; DB 6; Length 5111;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGTAAGTAAGTCTCAATTTGTTACTGAGAGTGGTGGCTGCTGGA 60
Db 470 ATGGCCCATGAGATGATGTAAGTAAGTCTCAATTTGTTACTGAGAGTGGTGGCTGCTGGA 529
QY 61 AGTGAAGGGAAGAAAGTGTGCTAATGTAAGCCGCGCAATTTGGAATTAACATATCC 120
Db 530 AGTGAAGGGAAGAAAGTGTGCTAATGTAAGCCGCGCAATTTGGAATTAACATATCC 589
QY 121 CACATTTTGAAGCCATTAATATCACTGCTCCAGCTTATGAGAGGAGTTGCAACAG 180
Db 590 CACATTTTGAAGCCATTAATATCACTGCTCCAGCTTATGAGAGGAGTTGCAACAG 649
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCAATTCAGGAAACATTAAGTTGACATTAT 240
Db 650 GACAAAGTGTAAATTAACAGAGCTCATCCAGCAATTCAGGAAACATTAAGTTGACATTAT 709
QY 241 TGAAGCAGAAAGTGTATTAACATCAAGAGTCCCAAGATGGTGGCTCTCTCTTCA 300
Db 710 TGAAGCAGAAAGTGTATTAACATCAAGAGTCCCAAGATGGTGGCTCTCTCTTCA 769
QY 301 GACTGTTTCTCACTGTAATTTCTGGGTAATGAGAGAGGTTCAACTCTGTCACCTG 360
Db 770 GACTGTTTCTCACTGTAATTTCTGGGTAATGAGAGAGGTTCAACTCTGTCACCTG 829
QY 361 CTTCAGAGTGGGTTTGTGAGATTTCTCTGTTGTTTCTCGGCTCTGTGAAGAAATCC 420
Db 830 CTTCAGAGTGGGTTTGTGAGATTTCTCTGTTGTTTCTCGGCTCTGTGAAGAAATCC 889
QY 421 ACTCTAGTCCCACTGATTTCTCAGAGCTTCTCAAGCTTCACTGTTGCAAGATTGGGCCAAC 480
Db 890 ACTCTAGTCCCACTGATTTCTCAGAGCTTCTCAGAGCTTCTCACTGTTGCAAGATTGGGCCAAC 949
QY 481 CGAATTTCTCCAAATCTTTATCTTGGCTGCGAGAGATGTCTCAACAAAGAGCTGATA 540

Db	950	CGAATCTTCCCACTTTATCTTGCGTCCAGGAGATGTCTCAACAAGAGCTATG	1009
Qy	541	CAGAGATGGATTTGGTATGTGTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT	600
Db	1010	CAGCAAAVGGATTTGGTATGTGTAAATGCCAGCAATACCTGTCCAAAGCTGACTTT	1069
Qy	601	ATCCCGAGTCTCAATTTCTGCTGCTGTGAATGACAGCTTTTGTGAATAATTTTG	660
Db	1070	ATCCCGAGTCTCAATTTCTGCTGCTGTGAATGACAGCTTTTGTGAATAATTTTG	1129
Qy	661	CCGTTGTTGACAAATCAGTATGATTTCAATGAGAAACAAAAGCTTCAATGATGTGT	720
Db	1130	CCGTTGTTGACAAATCAGTATGATTTCAATGAGAAACAAAAGCTTCAATGATGTGT	1189
Qy	721	CTAGTGCATGTTTAGCTGGGATCTCCGCTCCGCGCCACATCGTATTCGCTACATCATG	780
Db	1190	CTAGTGCATGTTTAGCTGGGATCTCCGCTCCGCGCCACATCGTATTCGCTACATCATG	1249
Qy	781	AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAGACCTACT	840
Db	1250	AAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAGACCTACT	1309
Qy	841	ATATCTCCAACTTCATTTTCTGGGCCAACTCCTGGAATATGAGAAAGATTAAAGAC	900
Db	1310	ATATCTCCAACTTCATTTTCTGGGCCAACTCCTGGAATATGAGAAAGATTAAAGAC	1369
Qy	901	CAGACTGAGCATCAGGGCCAAAGAGCAACTCAGCTGTGACCTCGAGAGAGCCCAAT	960
Db	1370	CAGACTGAGCATCAGGGCCAAAGAGCAACTCAGCTGTGACCTCGAGAGAGCCCAAT	1429
Qy	961	GAACTGTCTCTGCTGTCTCAGAGGGTGAACAGAAAAGCAGAGCCCTCAGTCCACCC	1020
Db	1430	GAACTGTCTCTGCTGTCTCAGAGGGTGAACAGAAAAGCAGAGCCCTCAGTCCACCC	1489
Qy	1021	TGAGCGGACCTCTCTCCTCCTAGAGGAGCAGAGCAAAAGCCCTGATCCCGCAGCTGT	1080
Db	1490	TGAGCGGACCTCTCTCCTCCTCCTCAGAGGAGCAGAGCAAAAGCCCTGATCCCGCAGCTGT	1549
Qy	1081	CCCAGGGTGGCCAGCGTGCAGCGCTGCCTGTGTAGAGACAGCCCGTGTACAGGGGCTC	1140
Db	1550	CCCAGGGTGGCCAGCGTGCAGCGCTGCCTGTGTAGAGACAGCCCGTGTACAGGGGCTC	1609
Qy	1141	AGTGGGCTGCACCTGTCCGAGACAGAGCTGGAAGACAGCAATTAAGCTCAAGGCTTCCT	1200
Db	1610	AGTGGGCTGCACCTGTCCGAGACAGAGCTGGAAGACAGCAATTAAGCTCAAGGCTTCCT	1669
Qy	1201	TCTCTGATATCAAAACAGTTTCATTTGAGCAGAGATGGAGCATCTCTTAATGAGCTTC	1260
Db	1670	TCTCTGATATCAAAACAGTTTCATTTGAGCAGAGATGGAGCATCTCTTAATGAGCTTC	1729
Qy	1261	TCTCTATCAGAAAGATGCTTTGAAATCTACAAACCTTCCACTACTCTGGATGGAGCAAC	1320
Db	1730	TCTCTATCAGAAAGATGCTTTGAAATCTACAAACCTTCCACTACTCTGGATGGAGCAAC	1789
Qy	1321	AAGCTATGCGAGTTCTCCCTGTTTGAAGAACTATGAGAGCAGACTTCCGAAAACAGTCT	1380
Db	1790	AAGCTATGCGAGTTCTCCCTGTTTGAAGAACTATGAGAGCAGACTTCCGAAAACAGTCT	1849
Qy	1381	GATTAAGAGAGAGCAGCATCCCAAGAGCTGCAGACCGCGAGCCCTTCAGACAGCCAG	1440
Db	1850	GATTAAGAGAGAGCAGCATCCCAAGAGCTGCAGACCGCGAGCCCTTCAGACAGCCAG	1909
Qy	1441	AGCAAGCATTTGCTGGTGCAGAACAGCAGCAGTGGCACCGCCAGAGGTCCCTTTTA	1500
Db	1910	AGCAAGCATTTGCTGGTGCAGAACAGCAGCAGTGGCACCGCCAGAGGTCCCTTTTA	1969
Qy	1501	TCTCCACGTCATCGAAGTGGAGGCTGAGAGCAATTAACAACACAGGTTCTTTTGGGC	1560
Db	1970	TCTCCACGTCATCGAAGTGGAGGCTGAGAGCAATTAACAACACAGGTTCTTTTGGGC	2029
Qy	1561	CTTTTCACACAGCAGACACACTCAGAAAGTGTGCTGGGCTTTAAAGGGCTGGCAC	1620
Db	2030	CTTTTCACACAGCAGACACACTCAGAAAGTGTGCTGGGCTTTAAAGGGCTGGCAC	2089

Qy	1621	TCGATATCTTGAGCCGCCAGACCTCTACCCCTTCTCTGACAGCAGCTGATTTTGGC	1680
Db	2090	TCGATATCTTGAGCCGCCAGACCTCTACCCCTTCTCTGACAGCAGCTGATTTTGGC	2149
Qy	1681	ACAGAGTCTTCACATTTCTACTCTGCTCAGCATCTACAGAGGAGTGGCAGTTACTCT	1740
Db	2150	ACAGAGTCTTCACATTTCTACTCTGCTCAGCATCTACAGAGGAGTGGCAGTTACTCT	2209
Qy	1741	GCCTACAGTGCAGCAGCTGCGCCACTTGCAGAGCAAGCTATCTGTGCGCAGGGCG	1800
Db	2210	GCCTACAGTGCAGCAGCTGCGCCACTTGCAGAGCAAGCTATCTGTGCGCAGGGCG	2269
Qy	1801	CAGAAAGCAAGTGCAGAGCTGACTCGCGCGAGCTGGCATGAAAGAGCCCTTTGAA	1860
Db	2270	CAGAAAGCAAGTGCAGAGCTGACTCGCGCGAGCTGGCATGAAAGAGCCCTTTGAA	2329
Qy	1861	AAGCAGTTTAAACGAGAAAGCTGCCAATTGSAAATTTGAGAGAGCATATGTCAAGAAC	1920
Db	2330	AAGCAGTTTAAACGAGAAAGCTGCCAATTGSAAATTTGAGAGAGCATATGTCAAGAAC	2389
Qy	1921	AGGTACCGGAAAGAGCTGGGAAAGTGGGAGTGCAGTACGTTTTCGGGAGCATGGA	1980
Db	2390	AGGTACCGGAAAGAGCTGGGAAAGTGGGAGTGCAGTACGTTTTCGGGAGCATGGA	2449
Qy	1981	ATCATTTAGTGTCTC	1995
Db	2450	ATCATTTAGTGTCTC	2464

RESULT 18
BC031643
LOCUS BC031643 2807 bp mRNA linear PRI 26-JUN-2002
DEFINITION Homo sapiens, clone IMAGE:5176724, mRNA, partial cds.
ACCESSION BC031643
VERSION BC031643.1 GI:21594972
KEYWORDS

SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2807)
JOURNAL Strausberg, R.

REMARK
COMMENT Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Guarante, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLNI at: <http://image.llnl.gov>
Series: IRAX Plate: 51 Row: n Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES
source

1..2807
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:5176724"
/tissue_type="Brain, lung, Testis, adult, pooled whole"
/clone_id="NIH_MGC_115"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
<1..1852
/codon_start=2
/product="unknown (protein for IMAGE:5176724)"
/protein_id="AAH31643.1"
/db_xref="GI:21594973"
/translation="CSKLMRRRLQODKVLITELIHSKAKHVIDSCOKVYVDOGSQ
DVASISDDCLFVLGLKESFNSVHLAAGFAFSPKCPGALCEGSTVPTCTISPC
LPVANGIPRILPNIYIGORDVANKEMOONGIYVLANSTCPKPDFPESHFLRV
PVNDSFEKILPMLDKSVDFIEKAKSNQCVLVHCLAGISRSATTAIYIMKMDSL
DEAYRFLVKEKRPITSFNFNLGQLLDIERKIKNOTAGSFGPSKULHLKEPVEPA
VSEGGQSETPPLPCDSATSEBAQORPHAPASVPSVPSVPSLLEDSLVQALSLG
HLSADRLDESNKLRSFSLDIKSVASASMAASLHGSSEDELEVYKSTLIDGINK
LCQSPVOELSEOTPEPSPDKEASIPKLOTAPSPDSQKRLSHVRSISGTAORSL
LSPLHSGSVEDNHTSFLRGLSTSOOHLTKSAGLIGKMSHSLIAPORTSPSTLSM
YRATESHPYSASAIYGGASYSAYSCSOLPTGDOYTVRRROKPSDRADSRSHWE
ESPFEKQFKRSCQMEGESIMSENSREELGRVGSQSSFSGSMEIIEVS"

CDS

ORIGIN

Query Match 85.0%; Score 1696; DB 9; Length 2807;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

147 CTGCTCCAGCTTATGAGGGAAGGTTGCAACAGCAAGCAAGTTAATTACAGAGCTCAT 206
1 CTGCTCCAGCTTATGAGGGAAGGTTGCAACAGCAAGCAAGTTAATTACAGAGCTCAT 60

207 CCGAGCTTACGCGAACAATAAGTTGACATTGATGACATCAGAGGTTGATTTACGA 266
61 CCGAGCTTACGCGAACAATAAGTTGACATTGATGACATCAGAGGTTGATTTACGA 120

267 TCAAGCTCCAGAGTGTGCTCTCTCTCTGACATGTTTCTCACTGATCTTGGG 326
121 TCAAGCTCCAGAGTGTGCTCTCTCTCTGACATGTTTCTCACTGATCTTGGG 180

327 TAACTGAGAGAGGCTTCAACTCTGTTCACTGCTTGCAGTGGGTTTGCAGATTCTC 386
181 TAACTGAGAGAGGCTTCAACTCTGTTCACTGCTTGCAGTGGGTTTGCAGATTCTC 240

387 TCGTTGTTTCCCTGCGCTCTGTGAGAGAAATTCACCTTAAGCTTACCTGCAATTCTCA 446
241 TCGTTGTTTCCCTGCGCTCTGTGAGAGAAATTCACCTTAAGCTTACCTGCAATTCTCA 300

447 GCCTTGCTTACCTGTTGCCAATTTGGGCCAACCCGGAATTTCTCCGAATCTTTATCTTGG 506
301 GCCTTGCTTACCTGTTGCCAATTTGGGCCAACCCGGAATTTCTCCGAATCTTTATCTTGG 360

507 CTGCGACGAGATGTTCTCTCAACAAGAGCTGATACAGACAGAAATGGATTGTTATGTTT 566
361 CTGCGACGAGATGTTCTCTCAACAAGAGCTGATACAGACAGAAATGGATTGTTATGTTT 420

567 AAATGACAGCTATACCTGTCCAAAGCCTGACTTTATCCCGAGTCTCATTTCTCCGCTGT 626
421 AAATGACAGCTATACCTGTCCAAAGCCTGACTTTATCCCGAGTCTCATTTCTCCGCTGT 480

627 GCCTGTGAATGACAGCTTTTGTGAGAAATTTTGGCGGTGTTGGAACAATCAGTAGATT 686
481 GCCTGTGAATGACAGCTTTTGTGAGAAATTTTGGCGGTGTTGGAACAATCAGTAGATT 540

687 CATTGAGAAAGCAAAAGCTTCAATGATGTTGTTCTAGTGACATGTTAGCTGGAGATCTC 746
541 CATTGAGAAAGCAAAAGCTTCAATGATGTTGTTCTAGTGACATGTTAGCTGGAGATCTC 600

747 CCGCTCGGCCACATGCTATGCTTACATCATGAGAGATGAGATGTCATTTTATGATGA 806
601 CCGCTCGGCCACATGCTATGCTTACATCATGAGAGATGAGATGTCATTTTATGATGA 660

807 AGCTTACAGATTTGTGAGAAAGAAAGCACTACTATATCTCCAAACTTCAATTTTCTGGG 866

661 AGCTTACAGATTTGTGAGAAAGAAAGCACTACTATATCTCCAAACTTCAATTTTCTGGG 720
867 CCAACTCTCTGAGCTATGAGAAAGATTAAAGACCAAGCTGAGAGCTCAGGGCCAAAGAG 926
721 CCAACTCTCTGAGCTATGAGAAAGATTAAAGACCAAGCTGAGAGCTCAGGGCCAAAGAG 780

927 CAAACTCAAGCTGCTCAGCTGAGAAAGCAATGAACTGTCCTGCTGTCTCAGAGGG 986
781 CAAACTCAAGCTGCTCAGCTGAGAAAGCAATGAACTGTCCTGCTGTCTCAGAGGG 840

987 TGGACGAAAGAGGAGACGCCCTCAGTCACCTGTGCGGACTGTCTAATCCTCAGAGGC 1046
841 TGGACGAAAGAGGAGACGCCCTCAGTCACCTGTGCGGACTGTCTAATCCTCAGAGGC 900

1047 AGCAGAGCAAAAGCCCGTGTATCCGCCACGCTGCGCCAGCTGCCAGAGCTGACGCCCTC 1106
901 AGCAGAGCAAAAGCCCGTGTATCCGCCACGCTGCGCCAGCTGCCAGAGCTGACGCCCTC 960

1107 GCTGTTAGAGGACAGCCCGCTGCTGACAGGCGCTCAGTGGGCTGACCTGTCCGACAGAG 1166
961 GCTGTTAGAGGACAGCCCGCTGCTGACAGGCGCTCAGTGGGCTGACCTGTCCGACAGAG 1020

1167 GCTGGAAGACAGCAATTAAGCTCAAGCGTTCCTTCTCTGTGATATCAAAATCAAGTTTCATA 1226
1021 GCTGGAAGACAGCAATTAAGCTCAAGCGTTCCTTCTCTGTGATATCAAAATCAAGTTTCATA 1080

1227 TTCAGCCAGAGTGGCAGCATCTTACATGAGCTTCTCTCTCATCAGAAATGCTTTGGAATA 1286
1081 TTCAGCCAGAGTGGCAGCATCTTACATGAGCTTCTCTCTCATCAGAAATGCTTTGGAATA 1140

1287 CTACAAACCTTCCACACTCTGGAATGGAGCAACAGACTATGCGAGTTGCTCCCTGTGTA 1346
1141 CTACAAACCTTCCACACTCTGGAATGGAGCAACAGACTATGCGAGTTGCTCCCTGTGTA 1200

1347 GGAATCTATGAGAGCAAGTCCCGAAACAGTCTCTGATTAAGAGAGAGCAAGATCCCCCA 1406
1201 GGAATCTATGAGAGCAAGTCCCGAAACAGTCTCTGATTAAGAGAGAGCAAGATCCCCCA 1260

1407 GAAAGTGCAGACCGGCGCTTACAGACAGCCAGACAGCAAGCTTATGCTGCTGAGAAC 1466
1261 GAAAGTGCAGACCGGCGCTTACAGACAGCCAGACAGCAAGCTTATGCTGCTGAGAAC 1320

1467 CAGCAGAGTGGGACCGGCCAGAGTCCCTTTATATCTCACTGCAATCGAATGGGAGCGT 1526
1321 CAGCAGAGTGGGACCGGCCAGAGTCCCTTTATATCTCACTGCAATCGAATGGGAGCGT 1380

1527 GGAGAGCAATTAACACACAGCAAGCTTCTTTTGGCTTTTCCACAGCAGCAGACAGCTTAC 1586
1381 GGAGAGCAATTAACACACAGCAAGCTTCTTTTGGCTTTTCCACAGCAGCAGACAGCTTAC 1440

1587 GAAAGTGTGCTGGCTGGGCTTAAAGGCTGAGCACTGGAATATCTTGAGCCGCCAGACTTC 1646
1441 GAAAGTGTGCTGGCTGGGCTTAAAGGCTGAGCACTGGAATATCTTGAGCCGCCAGACTTC 1500

1647 TACCCCTTCCCTGACAGCAGCTGTGATTTTGGCAAGAGTCTTCACTTCACTCTGCG 1706
1501 TACCCCTTCCCTGACAGCAGCTGTGATTTTGGCAAGAGTCTTCACTTCACTCTGCG 1560

1707 CTGAGCATCTGAGAGGAGAGTGCAGTTACTGTGCTCAGCTGAGAGCCAGCTGAGCCAG 1766
1561 CTGAGCATCTGAGAGGAGAGTGCAGTTACTGTGCTCAGCTGAGAGCCAGCTGAGCCAG 1620

1767 TTGCGAGAGCAAAAGTATTTGTGTGCGAGGCGGAGAAAGCCAAAGTGAAGAGCTGACTC 1826
1621 TTGCGAGAGCAAAAGTATTTGTGTGCGAGGCGGAGAAAGCCAAAGTGAAGAGCTGACTC 1680

1827 GCGGCGGAGCTGGCAGTGAAGAGAGGCCCTTTTGAAGAGCTTTAAACGACAGAGCTGGCCA 1886
1681 GCGGCGGAGCTGGCAGTGAAGAGAGGCCCTTTTGAAGAGCTTTAAACGACAGAGCTGGCCA 1740

1887 AATGAAATTTGAGAGAGCATCATGTCAAGAACAGGTCAAGGAGAGCTGGGGAAGT 1946

Db 1741 AATGGAATTGGAGAGACATCATGTCAAGAAACAGGTCAAGGAAAGCTGGGAAAGT 1800
QY 1947 GGGCAGTCAGTCTAGCTTTTGGGGACAGATGAAATCATTTAGGTCTCC 1995
Db 1801 GGGCAGTCAGTCTAGCTTTTGGGGACAGATGAAATCATTTAGGTCTCC 1849

RESULT 19
AX099933 2118 bp DNA linear PAT 02-APR-2001
LOCUS Sequence 15 from Patent WO0120004.
DEFINITION AX099933
ACCESSION AX099933.1 GI:13538943
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Yue, H., Tang, Y.T., Bandman, O., Hillman, J.L., Baughn, M.R.,
Azimzai, Y. and Lu, D.A.
TITLE Protein phosphatase and kinase proteins
JOURNAL Patent: WO 0120004-A 15 22-MAR-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..2118
source
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID NO: 1234795CB1"

ORIGIN
Query Match 79.8%; Score 1593; DB 6; Length 2118;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 ATGGCCCATGAGATGATGAACTCAATTTGTTACTGAGAGGTGTGCTGTCTGAGAA 60
Db 25 ATGGCCCATGAGATGATGAACTCAATTTGTTACTGAGAGGTGTGCTGTCTGAGAA 84
QY 61 AGTGGAAACGAAAAAGTGTCTTAATTGATAGCCGGCCATTTGTGGAATACATATCATC 120
Db 85 AGTGGAAACGAAAAAGTGTCTTAATTGATAGCCGGCCATTTGTGGAATACATATCATC 144
QY 121 CACATTTTGGAGCCATTATATCAACCTCTCCAGCTTATGAAGGAAAGTTGCAACAG 180
Db 145 CACATTTTGGAGCCATTATATCAACCTCTCCAGCTTATGAAGGAAAGTTGCAACAG 204
QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGATTCCAGCGAAACATTAAGTTGACATTCAT 240
Db 205 GACAAAGTGTATTAACAGAGCTCATCCAGATTCCAGCGAAACATTAAGTTGACATTCAT 264
QY 241 TGCAGTCAGAAAGTGTGATTTACGATCAAGAGTCTCCAAAGATTTGCTCTCTCTTCA 300
Db 265 TGCAGTCAGAAAGTGTGATTTACGATCAAGAGTCTCCAAAGATTTGCTCTCTCTTCA 324
QY 301 GACTGTTTCTCACTGTAATCTCGGGTAACTGGAAGAGCTTCAACTCTGTGACCTG 360
Db 325 GACTGTTTCTCACTGTAATCTCGGGTAACTGGAAGAGCTTCAACTCTGTGACCTG 384
QY 361 CTTCGAGGTGGTTTGTCTGAGTTCTCTGTTTTCCTGGCTCTGTGAAAGAAATCC 420
Db 385 CTTCGAGGTGGTTTGTCTGAGTTCTCTGTTTTCCTGGCTCTGTGAAAGAAATCC 444
QY 421 ACTCTAGTCCCTACCTGCACTTTCTCAGCTTCTTAACCTGTTGCCAACATTGGGCCAAC 480
Db 445 ACTCTAGTCCCTACCTGCACTTTCTCAGCTTCTTAACCTGTTGCCAACATTGGGCCAAC 504
QY 481 GCAATTTCTCCAACTTTATCTTGCTGCTGCCAGCGAGAGTGTCTCAACAGAGAGTGTATA 540
Db 505 GCAATTTCTCCAACTTTATCTTGCTGCTGCCAGCGAGAGTGTCTCAACAGAGAGTGTATA 564
QY 541 CAGCAGAAATGGAGTGTATGTATGTTAAATGACCACTATACCTGTCCAAGGCTGACTTT 600

Db 565 CAGCAGAAATGGAGTGTATGTATGTTAAATGACCACAATACCTGTCCAAGGCTGACTTT 624
QY 601 ATCCCCGAGTCTCATTTTCTGCTGTGCTGTGGAATGACAGCTTTTGTGAGAAATTTTG 660
Db 625 ATCCCCGAGTCTCATTTTCTGCTGTGCTGTGGAATGACAGCTTTTGTGAGAAATTTTG 684
QY 661 CCGTGGTTGGAACAAATCAGTAGATTTCATTTGGAAGAAAGAAAGCCCTCAATGGATGTGT 720
Db 685 CCGTGGTTGGAACAAATCAGTAGATTTCATTTGGAAGAAAGAAAGCCCTCAATGGATGTGT 744
QY 721 CTAGTGAATGTTTGTAGTGGATCTCCGCTCCGACCATCTGCTACATGATGATG 780
Db 745 CTAGTGAATGTTTGTAGTGGATCTCCGCTCCGACCATCTGCTACATGATGATG 804
QY 781 AAGAGATGAGCATGTCTTTAGATGAAGCTTACAGATTTTGTGAAAGAAAGAAAGCTTACT 840
Db 805 AAGAGATGAGCATGTCTTTAGATGAAGCTTACAGATTTTGTGAAAGAAAGAAAGCTTACT 864
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAATCTCTGACCTATGAGAGAGTTAAGAAC 900
Db 865 ATATCTCCAACTTCAATTTTCTGGGCCAATCTCTGACCTATGAGAGAGTTAAGAAC 924
QY 901 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTTGAGAGGCAAAAT 960
Db 925 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTTGAGAGGCAAAAT 984
QY 961 GAACCTGTCTCTGCTGTCTCAGAGGGTGTGAGAGAAAGCGAGAGCGCCCTCATGTCACCC 1020
Db 985 GAACCTGTCTCTGCTGTCTCAGAGGGTGTGAGAGAAAGCGAGAGCGCCCTCATGTCACCC 1044
QY 1021 TGTGCCGACTCTGCTACCTCAGAGGCAAGAGCAAGAAAGCCCGTGCATCCGCGCAGCGTG 1080
Db 1045 TGTGCCGACTCTGCTACCTCAGAGGCAAGAGCAAGAAAGCCCGTGCATCCGCGCAGCGTG 1104
QY 1081 CCCAGCGTCCCGAGCGTCAAGCCGTGCTGTAGAGACAGCCCGCTGTACAGGCGCTC 1140
Db 1105 CCCAGCGTCCCGAGCGTCAAGCCGTGCTGTAGAGACAGCCCGCTGTACAGGCGCTC 1164
QY 1141 AGTGGCTGCACTGTCTCCGAGACAGAGCTGTGAGAGACAGCAATTAAGCTTAAGCTTCTTC 1200
Db 1165 AGTGGCTGCACTGTCTCCGAGACAGAGCTGTGAGAGACAGCAATTAAGCTTCAAGCTTCTTC 1224
QY 1201 TCTCTGATTCAAATCAGTTTCAATATTCAGCCAGCATGAGAGATCCTTACATGGCTTC 1260
Db 1225 TCTCTGATTCAAATCAGTTTCAATATTCAGCCAGCATGAGAGATCCTTACATGGCTTC 1284
QY 1261 TCTCTATCAGAAAGATGCTTTGGAATATCAAACTTCCACTACTGTGATGGAGCCAAAC 1320
Db 1285 TCTCTATCAGAAAGATGCTTTGGAATATCAAACTTCCACTACTGTGATGGAGCCAAAC 1344
QY 1321 AAGCTATGCCAGTTCTCCCTGTTTCAAGAACTATTCGAGAGACTCCGAAACCAAGT--C 1378
Db 1345 AAGCTATGCCAGTTCTCCCTGTTTCAAGAACTATTCGAGAGACTCCGAAACCAAGTTC 1404
QY 1379 CTGATTAAGAGAGAGCCAGCATCCCAAGAGCTGAGACCGCAGAGCTTCAGACAGCC 1438
Db 1405 CTGATTAAGAGAGAGCCAGCATCCCAAGAGCTGAGACCGCAGAGCTTCAGACAGCC 1464
QY 1439 AAGAGCAAGCAATTCAGTTCTGTCAGAAACCAAGCAGATGGACCGCCAGAGGTCCTTT 1498
Db 1465 AAGAGCAAGCAATTCAGTTCTGTCAGAAACCAAGCAGATGGACCGCCAGAGGTCCTTT 1524
QY 1499 TATCTCAGCTCAGATGAAAGTGGAGCGTGGAGAGCAATTAACACACAGCTTCTTTTGG 1558
Db 1525 TATCTCAGCTCAGATGAAAGTGGAGCGTGGAGAGCAATTAACACACAGCTTCTTTTGG 1584
QY 1559 GCGTTTCCACAGCAGAGCAGACCTCAAGAGTGTGCTGAGCTGGG--CCTTAAGGGCTGG 1617
Db 1585 GCGTTTCCACAGCAGAGCAGACCTCAAGAGTGTGCTGAGCTGGGCTTTAAGGGCTGG 1644
QY 1618 CACTCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTT 1677

Db 1645 CACTCGAATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACCAGAGCTGGTATTTT 1704
QY 1678 GCCACAGAGTCTTCACTTCTTACTGTGCTTCCATCTACGAGGCAAGTGGCAGTTAC 1737
Db 1705 GCCACAGAGTCTTCACTTCTTACTGTGCTTCCATCTACGAGGCAAGTGGCAGTTAC 1764
QY 1738 TCTGCTTACAGCTGACAGCCAGCTGCCCATTTGGGAGACCAAGTCTATTCTGTGCGCAGG 1797
Db 1765 TCTGCTTACAGCTGACAGCCAGCTGCCCATTTGGGAGACCAAGTCTATTCTGTGCGCAGG 1824
QY 1798 CGGCAAGAGCCAAAGTACAGAGCTGAAGTGGGCGGAGCTGGCATGAGAGAGCCCTTT 1857
Db 1825 CGGCAAGAGCCAAAGTACAGAGCTGAAGTGGGCGGAGCTGGCATGAGAGAGCCCTTT 1884
QY 1858 GAAAAGCAGTTTAAAGCAGAGAGCTGCCAAATGGAAATTTGGAGAGACATCATGTCAGAG 1917
Db 1885 GAAAAGCAGTTTAAAGCAGAGAGCTGCCAAATGGAAATTTGGAGAGACATCATGTCAGAG 1944
QY 1918 AACAGGTCAAGGAGAGAGCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTCGGGACAGATG 1977
Db 1945 AACAGGTCAAGGAGAGAGCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTCGGGACAGATG 2004
QY 1978 GAAATCATTGAGGTCTCC 1995
Db 2005 GAAATCATTGAGGTCTCC 2022

Search completed: June 22, 2004, 14:05:22
Job time : 12411 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

Search completed: June 22, 2004, 16:13:28
Job time : 7677 secs

OM nucleic - nucleic search, using sw model

Run on: June 22, 2004, 09:21:35 ; Search time 7676 Seconds
(without alignments)

7761.210 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccatgagatgatgg.....tggaaatcatgaggtctcc 1995

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 1419

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database :

EST:*

- 1: em_estda:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_huv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_pbg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	ID	Description
No.						

No matches found						

This Page Blank (uspto)

Wed Jun 23 07:45:02 2004

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2004, 03:44:37 ; Search time 1140 Seconds
(without alignments)
7434.347 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532
Perfect score: 1995
Sequence: 1 atggcccatgagatgattgg.....tggaaatcattgaggtctcc 1995

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 212409041 residues

Word size : 1419

Total number of hits satisfying chosen parameters: 28

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database :

N_Geneseq_29Jan04:*
1: geneseqn1980a:*
2: geneseqn1990a:*
3: geneseqn2000a:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2003a:*
8: geneseqn2003b:*
9: geneseqn2003c:*
10: geneseqn2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1995	100.0	5450	6	ACC60559 Polynucle
2	1995	100.0	5450	6	ACC60572 Polynucle
3	1893	94.9	3059	6	AAS15768 Aas15768 cDNA sequ
4	1893	94.9	3496	6	ABK47596 Abk47596 cDNA enco
5	1893	94.9	3544	5	AAS14639 Aas14639 Human cDN
6	1893	94.9	3544	5	ABK49402 Abk49402 cDNA enco
7	1893	94.9	3766	6	ABK14474 Abk14474 Human pro
8	1893	94.9	4790	6	ABN83966 Abn83966 Human gen
9	1893	94.9	5145	5	ABV20833 Abv20833 Human pro
10	1893	94.9	5145	5	ABV21080 Abv21080 Human pro
11	1893	94.9	5145	5	ABV26680 Abv26680 Human pro
12	1893	94.9	5145	5	ABV20978 Abv20978 Human pro
13	1893	94.9	5145	5	ABV21092 Abv21092 Human pro
14	1893	94.9	5145	5	ABV21312 Abv21312 Human pro
15	1893	94.9	5145	5	ABV21316 Abv21316 Human pro
16	1893	94.9	5145	5	ABV26826 Abv26826 Human pro
17	1893	94.9	5145	5	ABV27131 Abv27131 Human pro
18	1893	94.9	5145	5	ABV26923 Abv26923 Human pro
19	1893	94.9	5145	5	ABV27135 Abv27135 Human pro
20	1893	94.9	5145	5	ABV28657 Abv28657 Human pro
21	1893	94.9	5145	5	ABV22827 Abv22827 Human pro
22	1893	94.9	5145	5	ABV26934 Abv26934 Human pro
23	1849	92.7	2966	4	AAB99685 Aab99685 Human pro

ALIGNMENTS

24	1842	92.3	2732	4	AAD09492	Aad09492 Human SGP
25	1842	92.3	3104	6	ABN59704	Abn59704 Novel hum
26	1791	89.8	2102	7	ADA53105	Ada53105 Human cod
27	1791	89.8	5111	6	ACC60521	Acc60521 Polynucle
28	1593	79.8	2118	4	AAF30479	Aaf30479 Human pro

RESULT 1
ACC60559
ID ACC60559 standard; cDNA; 5450 BP.

XX ACC60559;

DT 19-JUN-2003 (first entry)

DE Polynucleotide relating to the invention SEQ ID NO: 108.

XX Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic;
XX antipsoriatic; cardiant; cyostatic; gene therapy; liver disease;
XX proliferative disorder; renal failure; cardiovascular disorder;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.

OS Homo sapiens.

PN WQ020257460-A2.

PD 25-JUL-2002.

PF 20-DEC-2001; 2001WO-US050459.

PR 20-DEC-2000; 2000US-0256868P.

PR 30-MAR-2001; 2001US-0280186P.

PR 01-MAY-2001; 2001US-0287735P.

PR 05-JUN-2001; 2001US-0295848P.

PR 25-JUN-2001; 2001US-0300465P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;

XX Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Basolino D;

XX Kryslek S, Mcatee P, Suchard S, Banas D;

XX MPI; 2002-599721/64.

XX P-PsDB; ABR52381.

PT Novel polynucleotides encoding human phosphatase polypeptides useful in

PT the prevention or treatment of e.g. proliferative and cardiovascular

PT disorders.

XX Claim 1; Fig 13; 801pp; English.

CC The invention relates to a novel isolated nucleic acid comprising a

CC polynucleotide having a nucleotide sequence selected from 40

CC polynucleotides fully defined in the specification. The polynucleotide of

CC the invention has antiproliferative, hepatotropic, nephrotropic,

CC antarthritic, antipsoriatic, cardiant, and cyostatic activity. The

CC polynucleotide may have a use in gene therapy. A polynucleotide or

CC ameliorating a medical condition, e.g. a proliferative disorder. They are

CC also useful for treating e.g. liver disease, renal failure, immunological

CC disorders including arthritis and psoriasis, cardiovascular disorders

CC such as congenital heart defects and congestive heart failure, and

CC cancer. A method of the invention is useful for diagnosing a pathological

CC condition or susceptibility to a condition in a subject. The present

CC sequence is used in the exemplification of the invention

SO Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;

Query Match 100.0%; Score 1995; DB 6; Length 5450;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ATGGCCCATGAGATGATTTGAACTCAAAATTGTTACTGAGAGTGGTGGCTCTGCGAA 60
Db 538 ATGGCCCATGAGATGATTTGAACTCAAAATTGTTACTGAGAGTGGTGGCTCTGCGAA 597
OY 61 AGTGAACGGAAGAAAGTCTGCTAAATTGATAGCCGGCCCATTTGGGAAATCAATATCC 120
Db 598 AGTGAACGGAAGAAAGTCTGCTAAATTGATAGCCGGCCCATTTGGGAAATCAATATCC 657
OY 121 CACATTTTGAAGCCATTAAATATCAACTGCTCCAAAGCTTATGAAGCAAGTTGCAACAG 180
Db 658 CACATTTTGAAGCCATTAAATATCAACTGCTCCAAAGCTTATGAAGCAAGTTGCAACAG 717
OY 181 GACAAAGGTTAAATTAACAGAGCTATCCAGATTCAGGAAACATAAGTTGACATTTGAT 240
Db 718 GACAAAGGTTAAATTAACAGAGCTATCCAGATTCAGGAAACATAAGTTGACATTTGAT 777
OY 241 TGCAGTCAGAAAGTTAGTTAGTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 300
Db 778 TGCAGTCAGAAAGTTAGTTAGTTAGATCAAAAGCTCCCAAGATGTTGCTCTCTCTCA 837
OY 301 GACTGTTTCTCACTGTAATCTTGGGTAAACTGAGAAAGACTTCAACTCTGTTCACTTG 360
Db 838 GACTGTTTCTCACTGTAATCTTGGGTAAACTGAGAAAGACTTCAACTCTGTTCACTTG 897
OY 361 CTTCAGAGTGGGTTTGTGAGTTCTCGTGTGTTTCCCTGGCCCTTGTGAAAGAAATCC 420
Db 898 CTTCAGAGTGGGTTTGTGAGTTCTCGTGTGTTTCCCTGGCCCTTGTGAAAGAAATCC 957
OY 421 ACTCTAGTCCCTTACCTGCAATTTCTAGCCTTGTACTGTTGCCAAACTTGGGCCAAC 480
Db 958 ACTCTAGTCCCTTACCTGCAATTTCTAGCCTTGTACTGTTGCCAAACTTGGGCCAAC 1017
OY 481 CGAATCTTCCCAATCTTTATCTTGGCTGACAGAGATGTCCTCAACAGAGCTGATAT 540
Db 1018 CGAATCTTCCCAATCTTTATCTTGGCTGACAGAGATGTCCTCAACAGAGCTGATAT 1077
OY 541 CAGCAGAAATGGGATGTTATGTTAAATGACAGCTATACCTGTCCAAAGCTGACTTT 600
Db 1078 CAGCAGAAATGGGATGTTATGTTAAATGACAGCTATACCTGTCCAAAGCTGACTTT 1137
OY 601 ATCCCGAGTCTCATTTTCTGCGCTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
Db 1138 ATCCCGAGTCTCATTTTCTGCGCTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 1197
OY 661 CCGTGGTTGGAACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCCAAATGATGTGT 720
Db 1198 CCGTGGTTGGAACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCCAAATGATGTGT 1257
OY 721 CTAGTGCACTGTTTAGCTGGGATCTCCGCTCCGCAACCATGCTATGCGCTACATCATG 780
Db 1258 CTAGTGCACTGTTTAGCTGGGATCTCCGCTCCGCAACCATGCTATGCGCTACATCATG 1317
OY 781 AAGAGATGGAATGCTTTTAATGAAGTTTACATTTTGTGAAAGAAAGAAAGAAAGCTTACT 840
Db 1318 AAGAGATGGAATGCTTTTAATGAAGTTTACATTTTGTGAAAGAAAGAAAGAAAGCTTACT 1377
OY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCGGAATAAGAAAGAAAGTTTAAGAC 900
Db 1378 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCCGGAATAAGAAAGAAAGTTTAAGAC 1437
OY 901 CAGACTGAGATCAAGGCGCAAAAGCAAACTCAAGCTGCTGACCTTGGAGAGCCAAAT 960
Db 1438 CAGACTGAGATCAAGGCGCAAAAGCAAACTCAAGCTGCTGACCTTGGAGAGCCAAAT 1497
OY 961 GAACCTGTCCTGCTGCTCAAGAGGTGGAAGAAAGAGAGAGGCGCTCATGTCACCC 1020
Db 1498 GAACCTGTCCTGCTGCTCAAGAGGTGGAAGAAAGAGAGAGGCGCTCATGTCACCC 1557
OY 1021 TGTGCGAATCTGTAACCTCAGAGGCAAGAGCAAAAGCCGATGCTCCGCGCAGGCTG 1080
```

```
Db 1558 TGTGCGAATCTGTAACCTCAGAGGCAAGAGCAAAAGCCCGTGCATCCGCGCAGGCTG 1617
OY 1081 CCCAGGCTGCCAGCGTGCAGCGCTGCTGTTAAGAGACAGGCCCGTGGTTACAGGGCTTC 1140
Db 1618 CCCAGGCTGCCAGCGTGCAGCGCTGCTGTTAAGAGACAGGCCCGTGGTTACAGGGCTTC 1677
OY 1141 AGTGGCTGCACCTGTCCGCAAGACAGAGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
Db 1678 AGTGGCTGCACCTGTCCGCAAGACAGAGCTGGAAGACAGCAATTAAGCTCAAGGCTTCTTC 1737
OY 1201 TCTCTGGAATATCAATCAGTTTCAATTCAGGCACATGACAGCATCTTATATGGCTTC 1260
Db 1738 TCTCTGGAATATCAATCAGTTTCAATTCAGGCACATGACAGCATCTTATATGGCTTC 1797
OY 1261 TCTCTATCAGAAAGATGCTTTTGAATACTAACAACTTCCACTCACTCTGGAATGGGCAAC 1320
Db 1798 TCTCTATCAGAAAGATGCTTTTGAATACTAACAACTTCCACTCACTCTGGAATGGGCAAC 1857
OY 1321 AAGCTATGCAAGTTCTCCCTGTGTCAGGAACCTATCGAGACAGACTCCGAAACAGTCTCT 1380
Db 1858 AAGCTATGCAAGTTCTCCCTGTGTCAGGAACCTATCGAGACAGACTCCGAAACAGTCTCT 1917
OY 1381 GATTAAGAGGAAGCCAGCATCCCAAGAGCTGCAAGACCGGCAAGCTTCAAGACAGCCAG 1440
Db 1918 GATTAAGAGGAAGCCAGCATCCCAAGAGCTGCAAGACCGGCAAGCTTCAAGACAGCCAG 1977
OY 1441 AGCAAGCATTTGCAATTCGGTCAGAAACAGCAGCAGTGGCAACCGCCAGAGGTCCTTTTAA 1500
Db 1978 AGCAAGCATTTGCAATTCGGTCAGAAACAGCAGCAGTGGCAACCGCCAGAGGTCCTTTTAA 2037
OY 1501 TCTCCAGTCGATTCGAAGTGGGAGGCTGGAAGCAATTAACACACAGCTTCTTTTGGCC 1560
Db 2038 TCTCCAGTCGATTCGAAGTGGGAGGCTGGAAGCAATTAACACACAGCTTCTTTTGGCC 2097
OY 1561 CTTTCCACAGCAGCAGCAGCACTTCAAGAAATCTGCTGAGCTGAGGCTTAAAGGCTGACAC 1620
Db 2098 CTTTCCACAGCAGCAGCAGCACTTCAAGAAATCTGCTGAGCTGAGGCTTAAAGGCTGACAC 2157
OY 1621 TCGGATATCTTGGGCCCCAGACCTTCAACCTTTCCTGACCAAGCAGCTGATATTTTGGC 1680
Db 2158 TCGGATATCTTGGGCCCCAGACCTTCAACCTTTCCTGACCAAGCAGCTGATATTTTGGC 2217
OY 1681 ACAGAGTCTCAGACTTCTACTGCTGCTCAGCCATCTAAGGAGCAGTGCAGTACTCT 1740
Db 2218 ACAGAGTCTCAGACTTCTACTGCTGCTCAGCCATCTAAGGAGCAGTGCAGTACTCT 2277
OY 1741 GCCTACAGCTGCAGCAGCTGCCACTTGGCGAGAACCAAGTCTATTCTGCGCAGAGCGG 1800
Db 2278 GCCTACAGCTGCAGCAGCTGCCACTTGGCGAGAACCAAGTCTATTCTGCGCAGAGCGG 2337
OY 1801 CAGAAAGCAAGTACAGAGCTGACTGCGGCGGAGCTGSCATGAAGAGAGCCCTTTGAA 1860
Db 2338 CAGAAAGCAAGTACAGAGCTGACTGCGGCGGAGCTGSCATGAAGAGAGCCCTTTGAA 2397
OY 1861 AAGCAGTTTAAAGCAGAAAGCTGCCAAATGGAATTTGAGAGGCAATGTCATAGGAAC 1920
Db 2398 AAGCAGTTTAAAGCAGAAAGCTGCCAAATGGAATTTGAGAGGCAATGTCATAGGAAC 2457
OY 1921 AAGTCACGGAGAGAGCTGGGGAAGTGGGCAAGTCACTTAAGCTTTTGGGCGAGCATGGA 1980
Db 2458 AAGTCACGGAGAGAGCTGGGGAAGTGGGCAAGTCACTTAAGCTTTTGGGCGAGCATGGA 2517
OY 1981 ATCATTAAGGCTCC 1995
Db 2518 ATCATTAAGGCTCC 2532
```

RESULT 2
ACC60572
ID ACC60572 standard; cDNA; 5450 BP.
XX
ACC60572;
XX

DT 19-JUN-2003 (first entry)
XX Polynucleotide relating to the invention SEQ ID NO: 147.
DE
XX
XX Gene; *ss*; antiproliferative; hepatotropic; nephrotropic; antiarthritic;
XX antiproliferative; cardiatic; cytosolic; gene therapy; liver disease;
XX immunological disorder; renal failure; cardiovascular disorder;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
OS Homo sapiens.
XX
XX WO200257460-A2.
XX
XX 25-JUL-2002.
XX
XX 20-DEC-2001; 2001WO-US050459.
XX
XX 20-DEC-2000; 2000US-0256688P.
XX 30-MAR-2001; 2001US-0280186P.
XX 01-MAY-2001; 2001US-0287735P.
XX 05-JUN-2001; 2001US-0295848P.
XX 25-JUN-2001; 2001US-0300465P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Jackson DG, Feder J, Nelson T, Muntier G, Ramanathan C, Lee L,
XX Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
XX Krystek S, Mcatee P, Suchard S, Banas D;
XX
XX MPI: 2002-599721/64.
XX P-PSDB; ABR52407.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
XX the prevention or treatment of e.g. proliferative and cardiovascular
XX disorders.
XX
XX Example 7; Fig 19; 801pp; English.
XX
XX The invention relates to a novel isolated nucleic acid comprising a
XX polynucleotide having a nucleotide sequence selected from 40
XX polynucleotides fully defined in the specification. The polynucleotide of
XX the invention has antiproliferative, hepatotropic, nephrotropic,
XX antiarthritic, antiproliferative, cardiatic, and cytostatic activity. The
XX polynucleotide may have a use in gene therapy. A polynucleotide or
XX polypeptide of the invention is useful for preventing, treating or
XX ameliorating a medical condition, e.g. a proliferative disorder. They are
XX also useful for treating e.g. liver disease, renal failure, immunological
XX disorders including arthritis and psoriasis, cardiovascular disorders
XX such as congenital heart defects and congestive heart failure, and
XX cancer. A method of the invention is useful for diagnosing a pathological
XX condition or susceptibility to a condition in a subject. The present
XX sequence is used in the exemplification of the invention
XX
XX Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 1995; DB 6; Length 5450;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCCATGATGATTTGGAATCTCAATTTGTAAGAGCTTGTGGTCTGCTGGA 60
DB 538 ATGGCCCATGATGATTTGGAATCTCAATTTGTAAGAGCTTGTGGTCTGCTGGA 597
QY 61 AGTGAACGGAAGAGTGTCTAATGATAGCGGCATTTGTGAATCAATACATCC 120
DB 598 AGTGAACGGAAGAGTGTCTAATGATAGCGGCATTTGTGAATCAATACATCC 657
QY 121 CACATTTTGGAGCATTATATCAACTGCTCAAGCTTATGAAGCGAAGTTGCAACG 180
DB 658 CACATTTTGGAGCATTATATCAACTGCTCAAGCTTATGAAGCGAAGTTGCAACG 717
QY 181 GACAAAGTGTATTAACAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240

DB 718 GACAAAGTGTATTAACAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 777
QY 241 TGCAGTCAGAGGTTGTAGTTTACGATCAAGAGCTCCCAAGATGTTCTCTCTTCA 300
DB 778 TGCAGTCAGAGGTTGTAGTTTACGATCAAGAGCTCCCAAGATGTTCTCTCTTCA 837
QY 301 GACTGTTTCTACCTGTAATCTTGGGTAACTGAGAAAGAGCTTCAACTGTTCACCTG 360
DB 838 GACTGTTTCTACCTGTAATCTTGGGTAACTGAGAAAGAGCTTCAACTGTTCACCTG 897
QY 361 CTTCAGAGTGGTTTGTGAGTTCCTGCTGTTTCCCTGAGCTCTGTAAGAGAAATCC 420
DB 898 CTTCAGAGTGGTTTGTGAGTTCCTGCTGTTTCCCTGAGCTCTGTAAGAGAAATCC 957
QY 421 ACTTATGTCCTTACCTGATTTCTCAGCTTGTCTTACTTGTGCAACATTTGGCCACCC 480
DB 958 ACTTATGTCCTTACCTGATTTCTCAGCTTGTCTTACTTGTGCAACATTTGGCCACCC 1017
QY 481 CGAATTTCTCCCAATCTTTATATCTTGGCTGCGCAGAGAGTCTCTCAACAGAGCTGATA 540
DB 1018 CGAATTTCTCCCAATCTTTATATCTTGGCTGCGCAGAGAGTCTCTCAACAGAGCTGATA 1077
QY 541 CAGCAGATGGAGTTGTTATGTTAATGCAAGCTTATACCTGTCCAAAGCTGACTT 600
DB 1078 CAGCAGATGGAGTTGTTATGTTAATGCAAGCTTATACCTGTCCAAAGCTGACTT 1137
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 1138 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1197
QY 661 CCGTGTGTGACAAATCACTAGATTTTATTTGAGAAAGCAAAAGCTTCAATGATGTT 720
DB 1198 CCGTGTGTGACAAATCACTAGATTTTATTTGAGAAAGCAAAAGCTTCAATGATGTT 1257
QY 721 CTATGTCATGTTTATGCTGAGATCTCCGCTCCGACCATCGCTTATGCTTATCATCATG 780
DB 1258 CTATGTCATGTTTATGCTGAGATCTCCGCTCCGACCATCGCTTATGCTTATCATCATG 1317
QY 781 AAGAGATGAGATGCTTCTTATGATGAGCTTACAGATTTGTGAAAGAAAGAAAGCTTACT 840
DB 1318 AAGAGATGAGATGCTTCTTATGATGAGCTTACAGATTTGTGAAAGAAAGAAAGCTTACT 1377
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAGAC 900
DB 1378 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAGAC 1437
QY 901 CAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGACCTGAGAAAGCAAAAT 960
DB 1438 CAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGACCTGAGAAAGCAAAAT 1497
QY 961 GAACCTGTCCTGCTGCTCTCAGAGGCTGAGCAAGAAAGCGCCCTCACTAGTCCACCC 1020
DB 1498 GAACCTGTCCTGCTGCTCTCAGAGGCTGAGCAAGAAAGCGCCCTCACTAGTCCACCC 1557
QY 1021 TGTGCGCACTGTGCTACTCAAGAGGCTGAGCAAGAAAGCGCCCTCACTAGTCCACCC 1080
DB 1558 TGTGCGCACTGTGCTACTCAAGAGGCTGAGCAAGAAAGCGCCCTCACTAGTCCACCC 1617
QY 1081 CCCAGGTCGCGAGGCTGAGCGCTGCTGCTTATGAGGACAGCCCGCTGATCAAGGGCTC 1140
DB 1618 CCCAGGTCGCGAGGCTGAGCGCTGCTGCTTATGAGGACAGCCCGCTGATCAAGGGCTC 1677
QY 1141 AGTGGCTGCACTGTCTCCAGACAGGCTGAGCAAGCAAGCAATTAAGCTCAAGGTTCTTC 1200
DB 1678 AGTGGCTGCACTGTCTCCAGACAGGCTGAGCAAGCAAGCAATTAAGCTCAAGGTTCTTC 1737
QY 1201 TCTTGAATATCAATCACTTCAATTTTATGAGCAAGATGCGAGCATCTTACATGCTTC 1260
DB 1738 TCTTGAATATCAATCACTTCAATTTTATGAGCAAGATGCGAGCATCTTACATGCTTC 1797
QY 1261 TCTTGAATGAGATGCTTGAATTAATCAAACTTCACTTATGATGAGGAGCAAC 1320

Db 1798 TCCTCATCAGAAAGTGTGGAAATGACTACAAACCTTCCACTACTGTGATGGAGCAAC 1857
 QY 1321 AAGCTATGCCAGTCTCCCTGTTCCAGAACTATCGAGAGAGACTCCCGAAACCAAGTCTT 1380
 Db 1858 AAGCTATGCCAGTCTCCCTGTTCCAGAACTATCGAGAGAGACTCCCGAAACCAAGTCTT 1917
 QY 1381 GATTAAG 1440
 Db 1918 GATTAAG 1977
 QY 1441 AGCAAGGATTTGCTTTGGTTCAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 Db 1978 AGCAAGGATTTGCTTTGGTTCAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2037
 QY 1501 TCTCCACTGCATCGAAGTGGAG 1560
 Db 2038 TCTCCACTGCATCGAAGTGGAG 2097
 QY 1561 CTTTCCACAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
 Db 2098 CTTTCCACAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2157
 QY 1621 TCGGATATCTTTGGCCCCCGCCAGACCTCTACCCCTTCCCTGACCGAGAGCTGTATTTTGGCC 1680
 Db 2158 TCGGATATCTTTGGCCCCCGCCAGACCTCTACCCCTTCCCTGACCGAGAGCTGTATTTTGGCC 2217
 QY 1681 ACAGAGTCTCTACACTCTACTGCTCTGAGCCACTTACGAGAGAGAGAGAGAGAGAGAGAGAG 1740
 Db 2218 ACAGAGTCTCTACACTCTACTGCTCTGAGCCACTTACGAGAGAGAGAGAGAGAGAGAGAGAG 2277
 QY 1741 GCTTACAGCTGCGAGCCAGCTGCCCACTTGGGAGAGACCAAGTCTATTTCTGTGCGCAGAGCGG 1800
 Db 2278 GCTTACAGCTGCGAGCCAGCTGCCCACTTGGGAGAGACCAAGTCTATTTCTGTGCGCAGAGCGG 2337
 QY 1801 CAGAGGCAAGTACAGAGAGTGACTGCGCGCGGAGAGCTGCGCATTAAGAGAGAGAGAGAGAGAGAG 1860
 Db 2338 CAGAGGCAAGTACAGAGAGTGACTGCGCGCGGAGAGCTGCGCATTAAGAGAGAGAGAGAGAGAGAG 2397
 QY 1861 AAGCAGTTTAAACGAGAGAGCTGCCAATATGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
 Db 2398 AAGCAGTTTAAACGAGAGAGCTGCCAATATGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2457
 QY 1921 AGGTCAACGGAGAGAGCTGGGGGAAAGTGGGCAAGTCAAGTCTTTTGGGCAAGATGAA 1980
 Db 2458 AGGTCAACGGAGAGAGCTGGGGGAAAGTGGGCAAGTCAAGTCTTTTGGGCAAGATGAA 2517
 QY 1981 ATCATTGAGGTCCTCC 1995
 Db 2518 ATCATTGAGGTCCTCC 2532
 RESULT 3
 AAS15768
 ID AAS15768 standard; cDNA; 3059 BP.
 AC AAS15768;
 XX 18-JUN-2002 (first entry)
 XX cDNA sequence of human (dual specificity phosphatase) DUSP-10.
 DE
 XX
 XX Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke;
 KW neuronal degeneration syndrome; Alzheimer's disease; depression;
 KW schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
 KW osteoporosis; diabetes; gene therapy; chromosome 12; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 127..2124
 FT /*tag= a
 FT /product= "DUSP-10 protein"
 FT
 XX

PN WO200177340-A1.
 XX 18-OCT-2001.
 PD 06-APR-2001; 2001WO-EP003966.
 XX
 PF 10-APR-2000; 2000EP-00107143.
 XX
 PR (MERE) MERCK PATENT GMBH.
 XX
 PA Duecker K;
 XX
 PI WPI; 2002-010917/01.
 DR P-PSDB; A4U09946.
 DR
 PT Novel dual specificity phosphatase polypeptides useful for treating
 PT cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
 PT disease, depression, schizophrenia, asthma and immune disorders.
 XX
 PS Claim 5; Page 34-37; 43pp; English.
 CC The present invention relates to a new isolated dual specificity
 CC phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
 CC sequence that is fully defined in the specification. The invention also
 CC provides a sequence encoded by a 3059 nucleotide sequence fully defined
 CC in the specification, and a sequence having at least 95 % identity to the
 CC polypeptide, or fragments or variants of DUSP-10. The invention is useful
 CC for treating cancer e.g. leukemia, colon carcinoma, lung cancer,
 CC prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
 CC stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
 CC schizophrenia, cardiac myopathies, asthma, immune disorders,
 CC inflammatory processes e.g. arthritis, bowel disease, type I diabetes,
 CC osteoporosis, diabetes and diabetes associated diseases. The molecules of
 CC the invention are also useful as vaccines for inducing immunological
 CC response in a mammal, in disease diagnosis and in assays for screening
 CC agonistic or antagonistic compounds. Other uses of the invention include
 CC identifying membrane bound or soluble receptors, as a diagnostic reagent,
 CC in chromosome localisation studies, and as a valuable tool in tissue
 CC expression studies. The present sequence represents cDNA of the human
 CC dual specificity phosphatase, DUSP-10. DUSP-10 is located on chromosome
 CC 12
 XX
 SQ Sequence 3059 BP; 831 A; 761 C; 709 G; 758 T; 0 U; 0 Other;
 Query Match 94.9%; Score 1893; DB 6; Length 3059;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGGCCCATGAGATGATTTGGAACCAATTGTTACTGAGAGAGTTGGTGGCTGTGGAA 60
 Db 127 ATGGCCCATGAGATGATTTGGAACCAATTGTTACTGAGAGAGTTGGTGGCTGTGGAA 186
 QY 61 AGTGGAAAGGAAAAAGTGTGCTTAATTGATAGCCGGCATTGTGGAATACATATCC 120
 Db 187 AGTGGAAAGGAAAAAGTGTGCTTAATTGATAGCCGGCATTGTGGAATACATATCC 246
 QY 121 CACATTTTGAAGCATTAATATCACTGCTCCAAAGTTTGAAGCAAGTTGCAACAG 180
 Db 247 CACATTTTGAAGCATTAATATCACTGCTCCAAAGTTTGAAGCAAGTTGCAACAG 306
 QY 181 GACAAAGTGTAAATTACAGAGCTATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 240
 Db 307 GACAAAGTGTAAATTACAGAGCTATCCAGCATTCAGCGAAACATAAGTTGACATTGAT 366
 QY 241 TGCAAGTCAGAAAGTTGATTTACATGATCAAAAGTCCCAAGATGTGCTCTCTTTCA 300
 Db 367 TGCAAGTCAGAAAGTTGATTTACATGATCAAAAGTCCCAAGATGTGCTCTCTTTCA 426
 QY 301 GACTGTTTCTCACTGACTCTGGGGTAAACATGGAAGAGCTTCAACTGTTCACCTG 360
 Db 427 GACTGTTTCTCACTGACTCTGGGGTAAACATGGAAGAGCTTCAACTGTTCACCTG 486
 QY 361 CTTCAGAGTGGTGTGCTGAGTCTCTCGTTGTTCCCTGGCTCTGTGAAGGAAATCC 420

```

Db      487 CTTCGAGGAGGTTGCTGAGTTCTCTCGTTGTTCCCTGGCTCTGTGAAGAAATATC 546
Qy      421 ACTCTAGTCCCTACCTGCAATTTCTAGCTTGTCTTCTCTGTTCCAACTTGGCCAACTC 480
Db      547 ACTCTAGTCCCTACCTGCAATTTCTAGCTTGTCTTCTCTGTTCCAACTTGGCCAACTC 606
Qy      481 CGAATCTTCCCAATCTTATCTTGGCTGCAGCGAGATGCTCTCAACAGAGCTGATA 540
Db      607 CGAATCTTCCCAATCTTATCTTGGCTGCAGCGAGATGCTCTCAACAGAGCTGATA 666
Qy      541 CAGCAAAATGGGATTTGTTATGTTAAATGCGAGTATACCTGTCCAAAGCTGACTT 600
Db      667 CAGCAAAATGGGATTTGTTATGTTAAATGCGAGTATACCTGTCCAAAGCTGACTT 726
Qy      601 ATCCCGAGTCTCATTTCTCGCTGTGCTGTGATGACAGCTTTTGTGAATAATTTTG 660
Db      727 ATCCCGAGTCTCATTTCTCGCTGTGCTGTGATGACAGCTTTTGTGAATAATTTTG 786
Qy      661 CGGTGTTGGACAAATCAGTAGTTTCAATTGAGAAAGCAAAAGCTCCAAATGATGTT 720
Db      787 CGGTGTTGGACAAATCAGTAGTTTCAATTGAGAAAGCAAAAGCTCCAAATGATGTT 846
Qy      721 CTAGTCACTGTTTAACTGGGATCTCCGCTCGCCACCATCGCTATCGCTACATCATG 780
Db      847 CTAGTCACTGTTTAACTGGGATCTCCGCTCGCCACCATCGCTATCGCTACATCATG 906
Qy      781 AAGAGATGACATGTTCTTATGATGAACTTACAGATTTGTGAAGAAAAAAGACCTACT 840
Db      907 AAGAGATGACATGTTCTTATGATGAACTTACAGATTTGTGAAGAAAAAAGACCTACT 966
Qy      841 ATATCTCCAACTTAAATTTCTGGGCCAACTCTGTGACTATAGAGAAAGATTAGAAC 900
Db      967 ATATCTCCAACTTAAATTTCTGGGCCAACTCTGTGACTATAGAGAAAGATTAGAAC 1026
Qy      901 CAGACTGAGCATCAGGSCCAAGAGCAAACTCAAGCTGTGACCTGAGAGAACCAAT 960
Db      1027 CAGACTGAGCATCAGGSCCAAGAGCAAACTCAAGCTGTGACCTGAGAGAACCAAT 1086
Qy      961 GAACTGTCTCTGCTGTCTTCAAGGGGTGACAGAAAAAGGAGACGCTCTCACTCAACC 1020
Db      1087 GAACTGTCTCTGCTGTCTTCAAGGGGTGACAGAAAAAGGAGACGCTCTCACTCAACC 1146
Qy      1021 TGTGCGGATCTGTCTTACCTTCAAGGAGACAGAAAGGCTGTGATCCCGCACCGTG 1080
Db      1147 TGTGCGGATCTGTCTTACCTTCAAGGAGACAGAAAGGCTGTGATCCCGCACCGTG 1206
Qy      1081 CCCAGGTCGCCAGCGCTGACGCGTGTGATGAGGACAGCGCGTGTGATGAGGCTG 1140
Db      1207 CCCAGGTCGCCAGCGCTGACGCGTGTGATGAGGACAGCGCGTGTGATGAGGCTG 1266
Qy      1141 AGTGGGCTGACCTGTCCGAGACAGAGCTGGAAGACAGCAATAGTCAAGCGTTCTTC 1200
Db      1267 AGTGGGCTGACCTGTCCGAGACAGAGCTGGAAGACAGCAATAGTCAAGCGTTCTTC 1326
Qy      1201 TCTCTGGAATTCAAATCAGTTTATATTTAGCCAGCATGCGAGCATCTTACATGCTTC 1260
Db      1327 TCTCTGGAATTCAAATCAGTTTATATTTAGCCAGCATGCGAGCATCTTACATGCTTC 1386
Qy      1261 TCTCTCATGAAAGATGTTGAAATCTCAAACTTCCACTCTCTGATGGAGCAAC 1320
Db      1387 TCTCTCATGAAAGATGTTGAAATCTCAAACTTCCACTCTCTGATGGAGCAAC 1446
Qy      1321 AAGCTATGCGAGTTCTCCCTGTTCAGAACTATCGAGAGCATCCCGAAACAGTCTCT 1380
Db      1447 AAGCTATGCGAGTTCTCCCTGTTCAGAACTATCGAGAGCATCCCGAAACAGTCTCT 1506
Qy      1381 GATTAAGAGAAAGCAGCATCTCCCAAGAACTGCAACCGCGCTTCAAGACCGAG 1440
Db      1507 GATTAAGAGAAAGCAGCATCTCCCAAGAACTGCAACCGCGCTTCAAGACCGAG 1566
Qy      1441 AGCAAGCATTTGATTCGATCGATCAGAACAGAGAGTGGACCGCGCAGAGAGTCCCTTTTA 1500

```

```

Db      1567 AGCAAGCATTCATTCGCTCAGAAACAGACAGAGTGGACCGCCAGAGATCCCTTTTA 1626
Qy      1501 TCTCCAGTCATCGAAGTGGGAGCGTGGAGCAATTTACACACAGCTTCTCTTTGGC 1560
Db      1627 TCTCCAGTCATCGAAGTGGGAGCGTGGAGCAATTTACACACAGCTTCTCTTTGGC 1686
Qy      1561 CTTTCACAGCAGCAGACGCTTCAAGAGTCTGTGCGCTGGGCTTTAAGGCTGGCAC 1620
Db      1687 CTTTCACAGCAGCAGACGCTTCAAGAGTCTGTGCGCTGGGCTTTAAGGCTGGCAC 1746
Qy      1621 TCGGATATCTGGCCGCCGAGACTTACCCCTTCTCTGACAGACAGCTGTATTTTGGC 1680
Db      1747 TCGGATATCTGGCCGCCGAGACTTACCCCTTCTCTGACAGACAGCTGTATTTTGGC 1806
Qy      1681 ACAGAGTCTTCACTTCTACTCTGCTTACGATCTAGGAGGAGGAGGAGGAGTACTCT 1740
Db      1807 ACAGAGTCTTCACTTCTACTCTGCTTACGATCTAGGAGGAGGAGGAGGAGTACTCT 1866
Qy      1741 GCCTACAGCTGACGAGCTGCGCACTTGCAGAGACCAAGTCTATTCTGTGCGAGGCG 1800
Db      1867 GCCTACAGCTGACGAGCTGCGCACTTGCAGAGACCAAGTCTATTCTGTGCGAGGCG 1926
Qy      1801 CAGAAAGCAAGTACAGAGCTGACTGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAA 1860
Db      1927 CAGAAAGCAAGTACAGAGCTGACTGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAA 1986
Qy      1861 AAGCAGTTTAAACGCAAGAGCTGCCAATATGAAATTTGAGAGAGCATATGTCAGAAAC 1920
Db      1987 AAGCAGTTTAAACGCAAGAGCTGCCAATATGAAATTTGAGAGAGCATATGTCAGAAAC 2046
Qy      1921 AGGTACGAGGAAAGAGCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTGGGAGCATGAA 1980
Db      2047 AGGTACGAGGAAAGAGCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTGGGAGCATGAA 2106
Qy      1981 ATCATTGAGTCTCC 1995
Db      2107 ATCATTGAGTCTCC 2121

RESULT 4
ABK47596
ID ABK47596 standard; cDNA; 3496 BP.
XX
AC ABK47596;
XX
DT 02-JUL-2002 (first entry)
XX
DE cDNA encoding human dual-specificity phosphatase-3 (DSP-16) protein.
XX
KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW cancer; graft-versus-host disease; allergy; metabolic disease;
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
KW intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 562..2559
FT /tag=a
FT /product="human dual-specificity phosphatase-3 (DSP-16)
FT protein"
XX
XX MO200226997-A2.
XX
XX 04-APR-2002.
XX
XX 25-SEP-2001; 2001WO-US030124.
XX
XX 26-SEP-2000; 2000US-0235487P.
XX
XX (CEPT-) CEPTYR INC.

```

XX Luche RM, Wei B;
 XX WPI; 2002-315802/35.
 DR P-PSDB; RAU79156.
 XX
 FT New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 XX
 PS Claim 7; Fig 1; 87pp; English.
 XX
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity, for modulation of a proliferative response in a cell, survival
 CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present nucleic
 CC acid sequence is that of the human dual-specificity phosphatase-3 (DSP-
 CC 16) gene located on chromosome 12p. This sequence encodes the human DSP-
 CC 16 protein of the invention
 XX
 SQ Sequence 3496 BP; 939 A; 854 C; 856 G; 847 T; 0 U; 0 Other;
 Query Match 94.9%; Score 1893; DB 6; Length 3496;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCCCATGAGTATGTAAGTCACTCAATGTTTACTGAGAGGTGTGCTCTGCTGGA 60
 DB 562 ATGCCCATGAGTATGTAAGTCACTCAATGTTTACTGAGAGGTGTGCTCTGCTGGA 621
 QY 61 AGTGAACGGAAAAAGTGTGCTAATGTTAGCGCGCATTTGTGAAATACATCATCC 120
 DB 622 AGTGAACGGAAAAAGTGTGCTAATGTTAGCGCGCATTTGTGAAATACATCATCC 681
 QY 121 CACATTTTGAAGCCCTTAATATCACTGCTCCAGCTTATGAGGAGGTGGCAACG 180
 DB 682 CACATTTTGAAGCCCTTAATATCACTGCTCCAGCTTATGAGGAGGTGGCAACG 741
 QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGCACTTACGCAACATTAAGTTACATTGAT 240
 DB 742 GACAAAGTGTATTAACAGAGCTCATCCAGCACTTACGCAACATTAAGTTACATTGAT 801
 QY 241 TGCAGTCAGAGGTTGTATTAACAGTCAAGCTCCAGAGTGTGCTCTCTCTTCA 300
 DB 802 TGCAGTCAGAGGTTGTATTAACAGTCAAGCTCCAGAGTGTGCTCTCTCTTCA 861
 QY 301 GACTGTTTTCTACCTGTACTTCTGGGTAACTGGAAGAGCTTCAACTCTGTTCACTG 360
 DB 862 GACTGTTTTCTACCTGTACTTCTGGGTAACTGGAAGAGCTTCAACTCTGTTCACTG 921
 QY 361 CTGGAAGGTGGTGTGCTGAGTCTCTGTTGTTCCCGGCTCTGTGAGGAAATACC 420
 DB 922 CTGGAAGGTGGTGTGCTGAGTCTCTGTTGTTCCCGGCTCTGTGAGGAAATACC 981
 QY 421 ACTCTAGTCCCTACCTGCAATTTCTGAGCTTGCTTACCTGTGCAACATTTGGGCAAC 480
 DB 982 ACTCTAGTCCCTACCTGCAATTTCTGAGCTTGCTTACCTGTGCAACATTTGGGCAAC 1041
 QY 481 GCAATTTCTCCCAATCTTATCTTGGCTGCCAGAGATGTCCTCAACAGAGCTGATA 540
 DB 1042 GCAATTTCTCCCAATCTTATCTTGGCTGCCAGAGATGTCCTCAACAGAGCTGATA 1101
 QY 541 CAGCAGAAATGGAGTTGTTATGTGTTAAATGCAAGCTATACCTGCAAGCTGACCTT 600

DB 1102 CAGCAGAAATGGAGTTGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCTGACCTT 1161
 QY 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 660
 DB 1162 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTG 1221
 QY 661 CCGTGTGGAACAAATACAGTATATTTCAATGTGAGAAAGCAAAAGCTCCATGATGTGT 720
 DB 1222 CCGTGTGGAACAAATACAGTATATTTCAATGTGAGAAAGCAAAAGCTCCATGATGTGT 1281
 QY 721 CTAGTCACTGTTTGTAGTGGATCTCCGCTCCGACCATGCTATGCGCTACATCANG 780
 DB 1282 CTAGTCACTGTTTGTAGTGGATCTCCGCTCCGACCATGCTATGCGCTACATCANG 1341
 QY 781 AAGAGATGACATGCTTTAGATGATGACATTAAGATTTGTGAAAAGAAAGACTTACT 840
 DB 1342 AAGAGATGACATGCTTTAGATGATGACATTAAGATTTGTGAAAAGAAAGACTTACT 1401
 QY 841 ATATCTCAAACTTCAATTTTCTGGGCCAATCTCTGACCTATGAGAAAGATTAAGAAC 900
 DB 1402 ATATCTCAAACTTCAATTTTCTGGGCCAATCTCTGACCTATGAGAAAGATTAAGAAC 1461
 QY 901 CAGACTGAGATCAGGCGCAAGCAAACTCAAGCTGTGCACTGAGAGAGCAAT 960
 DB 1462 CAGACTGAGATCAGGCGCAAGCAAACTCAAGCTGTGCACTGAGAGAGCAAT 1521
 QY 961 GAACTGTCTCTGCTGTCTCAAGAGGTGACAGAAAAGAGAGAGCTTCAATCC 1020
 DB 1522 GAACTGTCTCTGCTGTCTCAAGAGGTGACAGAAAAGAGAGAGCTTCAATCC 1581
 QY 1021 TGTGCGCACTGTGCTACCTCAGAGCAGAGCAAGAAAGGCGCGTCAATCCCGCAGGCTG 1080
 DB 1582 TGTGCGCACTGTGCTACCTCAGAGCAGAGCAAGAAAGGCGCGTCAATCCCGCAGGCTG 1641
 QY 1081 CCAAGCTGCGCAGCGTGTGACAGCGCTGCTGTATGAGAGACAGCCGCTGTATCAGGCGCTC 1140
 DB 1642 CCAAGCTGCGCAGCGTGTGACAGCGCTGCTGTATGAGAGACAGCCGCTGTATCAGGCGCTC 1701
 QY 1141 AGTGGGCTGCACTGTGCGAGACAGAGCTGGAAGACGAAATTAAGCTCAAGCTTCTTC 1200
 DB 1702 AGTGGGCTGCACTGTGCGAGACAGAGCTGGAAGACGAAATTAAGCTCAAGCTTCTTC 1761
 QY 1201 TCTCTGATATCAAAATCAATAGTTTCAATATTCAGCCAGATGAGAGATCTTCAATGCTTC 1260
 DB 1762 TCTCTGATATCAAAATCAATAGTTTCAATATTCAGCCAGATGAGAGATCTTCAATGCTTC 1821
 QY 1261 TCTCTATCAGAAAGTCTTTGGAATCTACAAACCTTCACTACTGTGATGAGACCAAC 1320
 DB 1822 TCTCTATCAGAAAGTCTTTGGAATCTACAAACCTTCACTACTGTGATGAGACCAAC 1881
 QY 1321 AAGCTATGCAAGTCTCCCTGTTCAGGAATTAATGGAAGACCTCCGAAACCAATGCTT 1380
 DB 1882 AAGCTATGCAAGTCTCCCTGTTCAGGAATTAATGGAAGACCTCCGAAACCAATGCTT 1941
 QY 1381 GATTAAGAGAGAGCAGATCCCAAGAGCTCAGACGCGCAGGCTTCAAGACAGCAG 1440
 DB 1942 GATTAAGAGAGAGCAGATCCCAAGAGCTCAGACGCGCAGGCTTCAAGACAGCAG 2001
 QY 1441 AGCAAGGATGATTCATTCGCTCAGAACCAAGCAGAGTGGCAACGCGCCAGAGTCCCTTTA 1500
 DB 2002 AGCAAGGATGATTCATTCGCTCAGAACCAAGCAGAGTGGCAACGCGCCAGAGTCCCTTTA 2061
 QY 1501 TCTCACTGATGGAAGTGGAGCGTGGAGCAATTAACACACGAGCTTCTTTTCCGCG 1560
 DB 2062 TCTCACTGATGGAAGTGGAGCGTGGAGCAATTAACACACGAGCTTCTTTTCCGCG 2121
 QY 1561 CTTTCCACAGCAGCAGACACCTCAAGAGTGTGCTGCGCTTAAAGGGCTGGCAC 1620
 DB 2122 CTTTCCACAGCAGCAGACACCTCAAGAGTGTGCTGCGCTTAAAGGGCTGGCAC 2181
 QY 1621 TCGGATATCTTGGCCCCCAGACCTTATACCTTCCCTGACAGAGCTGTATTTTGGC 1680
 DB 2182 TCGGATATCTTGGCCCCCAGACCTTATACCTTCCCTGACAGAGCTGTATTTTGGC 2241

QY	1661	ACAGAGTCTCACACTTCTACTCTGGCCCTCAGCACTTACGGAGGAGGACGAGTTACTCT	1740
Db	2242	ACAGAGTCTCTCACACTTCTACTCTGGCCCTCAGCACTTACGGAGGAGGACGAGTTACTCT	2301
QY	1741	GCCTACAGCTCGAGCCAGCTGCCCCATTGCGAGACCAAGTCTATTCTGTGCGCAGGCGG	1800
Db	2302	GCCTACAGCTCGAGCCAGCTGCCCCATTGCGAGACCAAGTCTATTCTGTGCGCAGGCGG	2361
QY	1801	CAGAAAGCCAAAGTGAAGAGCTGACTCTGCGGCGGAGAGCTGGCATGAAGAGAGCCCTTTGAA	1860
Db	2362	CAGAAAGCCAAAGTGAAGAGCTGACTCTGCGGCGGAGAGCTGGCATGAAGAGAGCCCTTTGAA	2421
QY	1861	AAGCAGTTTAAACCGCAGAAGCTGCCAAATGGAATTGTGAGAGAGCATGTCAGAGAAC	1920
Db	2422	AAGCAGTTTAAACCGCAGAAGCTGCCAAATGGAATTGTGAGAGAGCATGTCAGAGAAC	2481
QY	1921	AGGTCAACGGGAAGAGCTGGGGGAAAGTGGGCACTGATCTTTCGGGAGCAGCATGGA	1980
Db	2482	AGGTCAACGGGAAGAGCTGGGGGAAAGTGGGCACTGATCTTTCGGGAGCAGCATGGA	2541
QY	1981	ATCATTTGAGGTCTCC	1995
Db	2542	ATCATTTGAGGTCTCC	2556
RESULT 5			
ID	AA514639	standard; cDNA, 3544 BP.	
XX	AA514639;		
AC			
XX			
DT	18-DEC-2001	(first entry)	
DE	Human cDNA encoding dual specificity phosphatase 21117.		
XX			
KW	Human; ss; dual specificity phosphatase 21117; hepatotropic; cytotostatic; hematopoietic disorder; autoimmune disorder; diabetes mellitus; rheumatoid arthritis; multiple sclerosis; Crohn's disease; liver disorder; erythroid associated disorder; haemolytic anaemia; cellular proliferative; differentiative disorder; leukaemia; acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer.		
KM			
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	589..2586	
FT		/*tag= a	
FT		/product= "Dual specificity phosphatase 21117"	
PN	WO200173059-A2.		
PD			
XX	04-OCT-2001.		
XX			
PF	23-MAR-2001; 2001WO-US009477.		
XX			
PR	24-MAR-2000; 2000US-0191858P.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Meyers RA;		
XX			
DR	WPI; 2001-611635/70.		
XX			
DR	P-PsDB; AAU09016.		
XX			
PT	New human dual specificity polypeptides and nucleic acids for diagnosis of disease and treatment of e.g. liver disorders.		
XX			
PS	Claim 1; Fig 1; 143pp; English.		
XX			
CC	The invention relates to two novel human dual specificity phosphatases designated 21117 and 38692, the nucleic acids encoding them (including fragments, allelic variants, their complements or nucleic acids that		
CC			

Query Match	94.9%	Score 1893;	DB 5;	Length 3544;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1993;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY 1	ATGGCCCATGATGATTTGAACTCAAAATTTTACTAGAGAGTGGTGGCTCTGCTGAA	60		
Db 589	ATGGCCCATGATGATTTGAACTCAAAATTTTACTAGAGAGTGGTGGCTCTGCTGAA	648		
QY 61	AGTGAACGGAAGAAAGTCTCTAATTAATGATGACGGCCATTGTGGAATACATATACCC	120		
Db 649	AGTGAACGGAAGAAAGTCTCTAATTAATGATGACGGCCATTGTGGAATACATATACCC	708		
QY 121	CACATTTTGGAGGCATTATATATCAACTGCTCCAACTTATGAAGCGAAGTTGCAACAG	180		
Db 709	CACATTTTGGAGGCATTATATATCAACTGCTCCAACTTATGAAGCGAAGTTGCAACAG	768		
QY 181	GACAAAGCTTAAATTAACAGAGCTCATCCAGATTCAGCGAACAATAAGTTGACATTGAT	240		
Db 769	GACAAAGCTTAAATTAACAGAGCTCATCCAGATTCAGCGAACAATAAGTTGACATTGAT	828		
QY 241	TGCAGTCGAAAGTGTGATGTTATGCATCAAAAGCTCCCAAGATGTGTCTCTCTCTCA	300		
Db 829	TGCAGTCGAAAGTGTGATGTTATGCATCAAAAGCTCCCAAGATGTGTCTCTCTCTCA	888		
QY 301	GACTGTTTTCACATGTACTTCTGGGGTAACTGAGAAAGAGCTTCAACTCTGTTCACTG	360		
Db 889	GACTGTTTTCACATGTACTTCTGGGGTAACTGAGAAAGAGCTTCAACTCTGTTCACTG	948		
QY 361	CTTGCAAGTGGGTTTCTGCAAGTCTCTCTGTTGTTTCTCTGGCCCTCTGTGAAGAAATCC	420		
Db 949	CTTGCAAGTGGGTTTCTGCAAGTCTCTCTGTTGTTTCTCTGGCCCTCTGTGAAGAAATCC	1008		
QY 421	ACTGATGCGCTTACTGATCTTCAGGCTTCGCTTACCTGTGTGCAACATTGGGGCAACC	480		
Db 1009	ACTGATGCGCTTACTGATCTTCAGGCTTCGCTTACCTGTGTGCAACATTGGGGCAACC	1068		
QY 481	CGAATTTCTTCCCAATCTTTATCTTGAGCTGCAGAGAGATGTCTTCAACAGAGCTGATA	540		
Db 1069	CGAATTTCTTCCCAATCTTTATCTTGAGCTGCAGAGAGATGTCTTCAACAGAGCTGATA	1128		
QY 541	CAGCAGAATGGAGTTGGTATGTGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT	600		
Db 1129	CAGCAGAATGGAGTTGGTATGTGTTAAATGCCAGCTATACCTGTCCAAAGCTGACTTT	1188		
QY 601	ATCCCGCAGTCTAATTTCTGCGGATGCTGTGTAATGACACTTTTGTGAAGAAATTTTG	660		
Db 1189	ATCCCGCAGTCTAATTTCTGCGGATGCTGTGTAATGACACTTTTGTGAAGAAATTTTG	1248		
QY 661	CCGTGTTTGGCAATCAGTATGATTTCAATTGAGAAAGCAAAAGCTCCCAATGATGTGTT	720		
Db 1249	CCGTGTTTGGCAATCAGTATGATTTCAATTGAGAAAGCAAAAGCTCCCAATGATGTGTT	1308		
QY 721	CTAGTGCATGTTTAGCTGGAGATCTCCCGCTCCGCACATCGCTATTCCTCAATCATG	780		

```
Db 1309 CTAGGCACTGTTAGCTGGAGTCTCCGCTCCGCAACATCGCTATCGCTTAATCATG 1368
Qy 781 AAGAGATGAGCACTCTCTTTAGATGAGACTTAACAGATTTTGAAAGAAAAAGCCTACT 840
Db 1369 AAGAGATGAGCACTCTTTAGATGAGACTTAACAGATTTTGAAAGAAAAAGCCTACT 1428
Qy 841 ATATCTCCAACTTCAATTTTCTGGGCACTCCGGACTATGAGAAAGAAATTAAAC 900
Db 1429 ATATCTCCAACTTCAATTTTCTGGGCACTCCGGACTATGAGAAAGAAATTAAAC 1488
Qy 901 CAGACTGAGCATGAGGCGCAAGAGCAAACTCAAGCTGTGCACTTGAGAAAGCCAAAT 960
Db 1489 CAGACTGAGCATGAGGCGCAAGAGCAAACTCAAGCTGTGCACTTGAGAAAGCCAAAT 1548
Qy 961 GAACCTGTCCCTGCTGTCTGAGAGGGTGGACAGAAAGCGAGAGCGCCCTCACTCCACC 1020
Db 1549 GAACCTGTCCCTGCTGTCTGAGAGGGTGGACAGAAAGCGAGAGCGCCCTCACTCCACC 1608
Qy 1021 TGTGCCGACTCTGTCTACCTGAGAGGCAAGCAAAAGCCCTGCACTCCGCGCAGCGTG 1080
Db 1609 TGTGCCGACTCTGTCTACCTGAGAGGCAAGCAAAAGCCCTGCACTCCGCGCAGCGTG 1668
Qy 1081 CCCAGCTGTCCGCTGAGAGCGCTGCTGTGAGAGACGCCCTGCTGTAAGAGGCGCTC 1140
Db 1669 CCCAGCTGTCCGCTGAGAGCGCTGCTGTGAGAGACGCCCTGCTGTAAGAGGCGCTC 1728
Qy 1141 AGTGGGCTGACCTGTCCGAGACAGGCTGAGAAACAGCAAAATTAAGCTCAAGCTTCTTC 1200
Db 1729 AGTGGGCTGACCTGTCCGAGACAGGCTGAGAAACAGCAAAATTAAGCTCAAGCTTCTTC 1788
Qy 1201 TCTCTGAGATCAAAATCAGTTTCATATTCAGCCAGATGAGCATCTTACATGAGCTTC 1260
Db 1789 TCTCTGAGATCAAAATCAGTTTCATATTCAGCCAGATGAGCATCTTACATGAGCTTC 1848
Qy 1261 TCCCTCATGAGAAATGTTTGGAAATCTAATAAATCTTCCATCTCTGAGTGGAGCAAC 1320
Db 1849 TCCCTCATGAGAAATGTTTGGAAATCTAATAAATCTTCCATCTCTGAGTGGAGCAAC 1908
Qy 1321 AAGCTATGCGAGTTCTCCCTGTTCCAGAACTATCGGAGAGACCTCCGAAACCAAGTCT 1380
Db 1909 AAGCTATGCGAGTTCTCCCTGTTCCAGAACTATCGGAGAGACCTCCGAAACCAAGTCT 1968
Qy 1381 GATTAAGAGAGAACCGACATCCCCAAAGAGCTGACAGCCGAGCTTCAAGACAGCAG 1440
Db 1969 GATTAAGAGAGAACCGACATCCCCAAAGAGCTGACAGCCGAGCTTCAAGACAGCAG 2028
Qy 1441 AGCAAGCAGATGATCGGTGAGAACAGACAGAGTGGACAGCCGAGAGGTCCTTTTAA 1500
Db 2029 AGCAAGCAGATGATCGGTGAGAACAGACAGAGTGGACAGCCGAGAGGTCCTTTTAA 2088
Qy 1501 TCTCCACTGCAATCGAAGTGGAGCGTGGAGCAATTAACAACACAGCTTCTTTTGGAG 1560
Db 2089 TCTCCACTGCAATCGAAGTGGAGCGTGGAGCAATTAACAACACAGCTTCTTTTGGAG 2148
Qy 1561 CTTTCCACAGCAGACAGACCTCAAGAGTCTGCTGGCTTGAAGGCTTGAGCAG 1620
Db 2149 CTTTCCACAGCAGACAGACCTCAAGAGTCTGCTGGCTTGAAGGCTTGAGCAG 2208
Qy 1621 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCAGACAGAGTGTGATTTTGGC 1680
Db 2209 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCAGACAGAGTGTGATTTTGGC 2268
Qy 1681 ACAGAGTCTCTACACTTCTACTCTGCTCAGCCATCTACGAGGAGGAGCCAGTTACTCT 1740
Db 2269 ACAGAGTCTCTACACTTCTACTCTGCTCAGCCATCTACGAGGAGGAGCCAGTTACTCT 2308
Qy 1741 GCTTACACTGAGCAGCAGCTGCCACTTGGCGAGACCAAGTCTATTCTGTGCGAGCGG 1800
Db 2329 GCTTACACTGAGCAGCAGCTGCCACTTGGCGAGACCAAGTCTATTCTGTGCGAGCGG 2388
Qy 1801 CAGAGCCAGATGACAGAGTGAATCTGGGCGAGCTGGCATGAAAGAGAGCCCTTTGAA 1860
```

```
Db 2389 CAGAGCCAGATGACAGACTTACTTCCGCGGAGCTGGCATGAAAGAGCCCTTTGAA 2448
Qy 1861 AAGCACTTTAAACGAGAGCTGCCAAATGGAATTTGGAGAGGATCATGTCAAGAGAAC 1920
Db 2449 AAGCACTTTAAACGAGAGCTGCCAAATGGAATTTGGAGAGGATCATGTCAAGAGAAC 2508
Qy 1921 AGGTCAAGGAGAGAGCTGGGAAAGTGGGAGTCAAGTCTTCTTGGGAGCATGGA 1980
Db 2509 AGGTCAAGGAGAGAGCTGGGAAAGTGGGAGTCAAGTCTTCTTGGGAGCATGGA 2568
Qy 1981 ATCATTGAGGCTCC 1995
Db 2569 ATCATTGAGGCTCC 2583

RESULT 6
ABK49402
ID ABK49402 standard; cDNA; 3544 BP.
XX
AC ABK49402;
XX
DT 02-JUL-2002 (first entry)
XX
DE cDNA encoding human dual specificity phosphatase 21117 protein.
XX
KW Human; dual specificity phosphatase 21117; erythroid-related disorder;
KW haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
KW erythrocytosis; liver-related disorder; cancer; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 589..2586
FT /tag= a
FT /product= "Human dual specificity phosphatase 21117"
FT /note= "Specifically claimed in claim 2"
XX
US2002034807-A1.
XX
PD 21-MAR-2002.
XX
PF 23-MAR-2001; 2001US-00816494.
XX
PR 24-MAR-2000; 2000US-0191858P.
XX
PA (MEYE/) MEYERS R A.
XX
PI Meyers RA;
XX
DR WPI; 2002-351088/38.
XX
P-PSDB; AAU79929.
XX
PT New nucleic acids, designated 36692 and 21117, encoding dual specificity
PT phosphatases for treating cell proliferation and differentiation
PT disorders including hematopoietic and erythroid-related disorders and
PT cancers.
XX
PS Claim 2; Fig 1; 76pp; English.
XX
CC The present invention relates to new nucleic acids designated 36692 and
CC 21117 encoding dual specificity phosphatase family members. The nucleic
CC acid, polypeptide encoded by it, and antibody specific for the
CC polypeptide may be used to diagnose and treat hematopoietic-related
CC disorders such as leukaemias and autoimmune diseases, erythroid-related
CC disorders such as anemias and erythrocytosis, liver-related disorders,
CC and cancers, particularly of the breast, colon, adipose, prostate and
CC lung. The present nucleic acid sequence encodes the human dual
CC specificity phosphatase 21117 protein of the invention, as described
CC above
XX
SQ Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;
Query Match 94.9%; Score 1893; DB 6; Length 3544;
```

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGCCCATGAGATGATGGAACTCAATATTGTAAGAGAGTTGGTGGCTCTGCTGGAA 60
DB ATGGCCCATGAGATGATGGAACTCAATATTGTAAGAGTTGGTGGCTCTGCTGGAA 648
QY 61 AGTGAACGGAAGAGTGTCTAATTGATAGCCGGCATTTGGATATCAATATACATCC 120
DB AGTGAACGGAAGAGTGTCTAATTGATAGCCGGCATTTGGATATCAATATACATCC 708
QY 121 CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAAGGAGTTGCAACAG 180
DB CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAAGGAGTTGCAACAG 768
QY 181 GACAAAGTGTATTATACAGAGCTCATTCAGCATTCAGCGAAATTAAGTTGACATTGAT 240
DB GACAAAGTGTATTATACAGAGCTCATTCAGCATTCAGCGAAATTAAGTTGACATTGAT 828
QY 241 TGCAGTCAGAAAGTTGATTAAGATCAAGATCCCAAGATGTTGGCTCTCTCTTCA 300
DB TGCAGTCAGAAAGTTGATTAAGATCAAGATCCCAAGATGTTGGCTCTCTCTTCA 888
QY 301 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGAGCTTCACTCTGTTCACTG 360
DB GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGAGCTTCACTCTGTTCACTG 948
QY 361 CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAGAGAAATCC 420
DB CTTCGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAGAGAAATCC 1008
QY 421 ACTCTAGTCCCTTACCTGCAATTTCTCAGCCCTTGTACTTCTGCAACATTTGGCCCAAC 480
DB ACTCTAGTCCCTTACCTGCAATTTCTCAGCCCTTGTACTTCTGCAACATTTGGCCCAAC 1068
QY 481 CGAATTTCTCCCAATCTTTATCTTGGGTGCGCAGCAGATGCTCTCAACAGAGCTGATA 540
DB CGAATTTCTCCCAATCTTTATCTTGGGTGCGCAGCAGATGCTCTCAACAGAGCTGATA 1128
QY 541 CAGCAGATGGAGTTGTTATGTTTAAATGCAAGCTATACCTGTCCAAAGCCTGACTTT 600
DB CAGCAGATGGAGTTGTTATGTTTAAATGCAAGCTATACCTGTCCAAAGCCTGACTTT 1188
QY 601 ATCCCGAGTCTATTTCCGCGTGGCCGTGTAATGACAGCTTTTGTGAGAAATTTTG 660
DB ATCCCGAGTCTATTTCCGCGTGGCCGTGTAATGACAGCTTTTGTGAGAAATTTTG 1248
QY 661 CCGTGGTTGGAACAATCAGTAGATTTCAATTGAGAAAGCAAAAGCCTCCAAATGATGTT 720
DB CCGTGGTTGGAACAATCAGTAGATTTCAATTGAGAAAGCAAAAGCCTCCAAATGATGTT 1308
QY 721 CTAGTGCATCTTTAGCTGGGATCTCCGCTCGGCCAATGCTATGCTTACATCATG 780
DB CTAGTGCATCTTTAGCTGGGATCTCCGCTCGGCCAATGCTATGCTTACATCATG 1368
QY 781 AAGAGATGACATGTTTATGATGAGCTTACAGATTTTGTGAGAAAGAAAGAAAGCTTACT 840
DB AAGAGATGACATGTTTATGATGAGCTTACAGATTTTGTGAGAAAGAAAGAAAGCTTACT 1428
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTTAAGAAC 900
DB ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTTAAGAAC 1488
QY 1429 ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTTAAGAAC 1488
DB ATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGACATATGAGAAAGATTTAAGAAC 1548
QY 901 CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAAGCTGTGTGACCTGAGAAAGCCAAAT 960
DB CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAAGCTGTGTGACCTGAGAAAGCCAAAT 1548
QY 961 GAACTGTCTCTGTCTCAGAGGTTGACAGAAAGGAGAGAGCCCTCAGTCCACCC 1020
DB GAACTGTCTCTGTCTCAGAGGTTGACAGAAAGGAGAGAGCCCTCAGTCCACCC 1548
QY 1021 TGTGCCGACTCTGTACTCTCAGAGGAGAGCAAAAGCCGCTGATCCCGCAGCGTG 1080
DB TGTGCCGACTCTGTACTCTCAGAGGAGAGCAAAAGCCGCTGATCCCGCAGCGTG 1080
```

```
DB 1609 TGTGCCGACTCTGTACTCTCAGAGGAGAGCAAAAGCCGCTGATCCCGCAGCGTG 1668
QY 1081 CCCAGCTGCCAGGCTGAGCGTGTGCTGTAGAGAGCAGGCCGCTGATCAGGGCTC 1140
DB CCCAGCTGCCAGGCTGAGCGTGTGCTGTAGAGAGCAGGCCGCTGATCAGGGCTC 1728
QY 1141 AGTGGGCTCAGCTGTCCGAGACAGAGCTGAGAGACAGCAATTAAGCTCAAGCTTCTTC 1200
DB AGTGGGCTCAGCTGTCCGAGACAGAGCTGAGAGACAGCAATTAAGCTCAAGCTTCTTC 1788
QY 1201 TCTTGGATATCAATCACTTATATTCAGCCAGATGCGAGCATCTTACATGCTTC 1260
DB TCTTGGATATCAATCACTTATATTCAGCCAGATGCGAGCATCTTACATGCTTC 1848
QY 1261 TCTCTACAGAAAGTCTTTGGAATCTAACAACTTCCACTCTGTGATGGAGCAAC 1320
DB TCTCTACAGAAAGTCTTTGGAATCTAACAACTTCCACTCTGTGATGGAGCAAC 1908
QY 1321 AAGCTATGCAAGTCTCCCTGTTGAGAACTATCGAGACAGACTCCGAAACCAAGTCT 1380
DB AAGCTATGCAAGTCTCCCTGTTGAGAACTATCGAGACAGACTCCGAAACCAAGTCT 1968
QY 1381 GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGACAGACCGCCAGGCTTCAAGACCCAG 1440
DB GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGACAGACCGCCAGGCTTCAAGACCCAG 2028
QY 1441 AGCAAGCATTTGATTTGGTCAAGAACAGACAGATGAGACCGCCAGAGTCCCTTTA 1500
DB AGCAAGCATTTGATTTGGTCAAGAACAGACAGATGAGACCGCCAGAGTCCCTTTA 2088
QY 1501 TCTCCACTGATCAAGAGTGGAGCGGTGAGAGCAATTTACACACAGCTTCTTTCGCG 1560
DB TCTCCACTGATCAAGAGTGGAGCGGTGAGAGCAATTTACACACAGCTTCTTTCGCG 2148
QY 1561 CTTCACACAGCCAGCAGCAGCACTCAAGAGTCTGTGGCTGGGCTTAAAGGCTGGCAC 1620
DB CTTCACACAGCCAGCAGCAGCACTCAAGAGTCTGTGGCTGGGCTTAAAGGCTGGCAC 2208
QY 1621 TGGGATATTTGGCCCCCAGACCTTACCCCTTCCCTAAGCAGAGCTGTATTTTGGC 1680
DB TGGGATATTTGGCCCCCAGACCTTACCCCTTCCCTAAGCAGAGCTGTATTTTGGC 2268
QY 1681 ACAGAGTCTCAGACTTCTACTCTGCTCAGCAGCATTAAGAGGAGAGTCCAGTACTCT 1740
DB ACAGAGTCTCAGACTTCTACTCTGCTCAGCAGCATTAAGAGGAGAGTCCAGTACTCT 2328
QY 1741 GCCTACAGCTGACAGCAGCTGCCACTTGCAGAGACCAAGTCTATTCTGTGCGCAGCGG 1800
DB GCCTACAGCTGACAGCAGCTGCCACTTGCAGAGACCAAGTCTATTCTGTGCGCAGCGG 2388
QY 1801 CAGAGCCAAAGTACAGAGCTGACTGTGCGCGGAGCTGGCATGAGAGAGCCCTTTGAA 1860
DB CAGAGCCAAAGTACAGAGCTGACTGTGCGCGGAGCTGGCATGAGAGAGCCCTTTGAA 2448
QY 1861 AAGCAGTTTAAACGAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTGAGAGAAC 1920
DB AAGCAGTTTAAACGAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTGAGAGAAC 2508
QY 1921 AGGTCAAGGAAAGAGCTGGGGAAGTGGGAGTCACTTATGCTTTTGGGCAAGATGAA 1980
DB AGGTCAAGGAAAGAGCTGGGGAAGTGGGAGTCACTTATGCTTTTGGGCAAGATGAA 2568
QY 1981 ATCATTTAGAGTCTCC 1995
DB ATCATTTAGAGTCTCC 2583
```

RESULT 7
ABK14474
ID ABK14474 standard; cDNA; 3766 BP.
XX
AC ABK14474;

DT	08-MAY-2002	(first entry)
XX		
DE	Human protein phosphatase 7 (PP7) CDNA sequence.	
XX		
KW	Human; protein phosphatase 7; PP7; immune system disorder; AIDS;	
KW	acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;	
KW	Crohn's disease; neurologic disorder; epilepsy; Huntington's disease;	
KW	dementia; Parkinson's disease; developmental disorder; Down's syndrome;	
KW	cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;	
KW	melanoma; myeloma sarcoma; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	538..2535
FT		/product= "Protein_phosphatase_7_(PP7)"
XX		
PN	WO200210363-A2.	
XX		
PD	07-FEB-2002.	
XX		
PF	26-JUL-2001; 2001WO-US023716.	
XX		
PR	28-JUL-2000; 2000US-0221679P.	
PR	03-AUG-2000; 2000US-022372P.	
PR	10-AUG-2000; 2000US-0224309P.	
PR	18-AUG-2000; 2000US-0226728P.	
PR	30-AUG-2000; 2000US-0229254P.	
PR	08-SEP-2000; 2000US-0231366P.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
PI	Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;	
PI	Stewart EA, Gandhi AR, Patterson C, Lee EA, Hatalla AW, Lu DAM;	
PI	Triboley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;	
PI	Walia NK, Kearney L;	
XX		
DR	WPI: 2002-188735/24.	
XX	P-PSDB; AAU75789.	
PT	New protein phosphatases, useful for diagnosing, treating or preventing	
PT	immune system disorders (e.g. Crohn's disease), neurological disorders	
PT	(e.g. Parkinson's disease), or cell proliferative disorders (e.g.	
PT	cancers).	
XX		
XX	Claim 5; Page 114-115; 117pp; English.	
XX		
CC	The present invention relates to a new polypeptide, a naturally occurring	
CC	amino acid sequence at least 95 % identical to it, a biologically active	
CC	fragment of it or an immunogenic fragment of it. The polypeptides,	
CC	polynucleotides, agonists and antagonists are useful for diagnosing,	
CC	treating or preventing disorders associated with aberrant expression of	
CC	protein phosphatases (PP), particularly immune system disorders e.g.	
CC	acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,	
CC	asthma or Crohn's disease, neurological disorders e.g. epilepsy.	
CC	Huntington's disease, dementia or Parkinson's disease, developmental	
CC	disorders e.g. Down's syndrome, or cell proliferative disorders e.g.	
CC	cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma	
CC	or sarcoma. The present nucleic acid sequence encodes human protein	
CC	phosphatase 7 (PP7) which is one of several human protein phosphatases	
CC	(AAU75783-AAU75792) of the invention	
XX		
SO	Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;	
Query Match	94.9%; Score 1893; DB 6; Length 3766;	
Best Local Similarity	99.9%; Pred. No. 0;	
Matches 1993; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
1	ATGGCCCATGGAGTATGGAACTCAATTGTACTGAGAGGTGTGGCTCTGCTGAA	60
538	ATGGCCCATGGAGTATGGAACTCAATTGTACTGAGAGGTGTGGCTCTGCTGAA	597

QY	61	AGTGGACGGAAAAAGTGGCTGTAATTGATAGCGGCGCATTTGTGGAATACATACATCC	120
Db	598	AGTGGAACGGAAAAAGTGGCTGCTAAATTGATAGCGGCGCATTTGTGGAATACATACATCC	657
QY	121	CACATTTTGGAAAGCCATTATATATCAACCTGCTCCAGCTTATGAAAGGAAGTTGCAACG	180
Db	658	CACATTTTGGAAAGCCATTATATATCAACCTGCTCCAGCTTATGAAAGGAAGTTGCAACG	717
QY	181	GACAAAGTGTAAATTACAGAGCTCATCCAGCAATTCAGCGAAACATTAAGGTTGCATTTGAT	240
Db	718	GACAAAGTGTAAATTACAGAGCTCATCCAGCAATTCAGCGAAACATTAAGGTTGCATTTGAT	777
QY	241	TGCAGTCAGAAAGTGTAGTTATTCAGATCAAAAGCTCCAAAGATGTGCTCTCTCTTCA	300
Db	778	TGCAGTCAGAAAGTGTAGTTATTCAGATCAAAAGCTCCAAAGATGTGCTCTCTCTTCA	837
QY	301	GACTGTTTTCTACCTGTAATTTCTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACCTG	360
Db	838	GACTGTTTTCTACCTGTAATTTCTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACCTG	897
QY	361	CTTGCAGGTGGGTTTGCTGAGTTCTCTCGTTGTTTCCCTGGGCGCTCGTGAAGAAAAATCC	420
Db	898	CTTGCAGGTGGGTTTGCTGAGTTCTCTCGTTGTTTCCCTGGGCGCTCTGTGAAGAAAAATCC	957
QY	421	ACTCTAGTCCCTACCTGCATTTTCTCAGCGCTTGCTTACCTGTGTCCAACTTGGGCCAAC	480
Db	958	ACTCTAGTCCCTACCTGCATTTTCTCAGCGCTTGCTTACCTGTGTGTCCAACTTGGGCCAAC	1011
QY	481	CGAATTTCTTCCCAATCTTTATCTTGTGGCTGGCGGAGATGTCCTCAACAAAGAGCTGATTA	540
Db	1018	CGAATTTCTTCCCAATCTTTATCTTGTGGCTGGCGGAGATGTCCTCAACAAAGAGCTGATTA	1077
QY	541	CAGCAGAAATGGAGTATGTTATGTGTTAAATGCGACGTAATACCTGTGCCAAAGCCTGACTTT	600
Db	1078	CAGCAGAAATGGAGTATGTTATGTGTTAAATGCGACGTAATACCTGTGCCAAAGCCTGACTTT	1137
QY	601	ATCCCCGAGTCTCATTTCTCGCTGCTGCTGTGAATGACAGCTTTTGTGAGAAAAATTTTG	660
Db	1138	ATCCCCGAGTCTCATTTCTCGCTGCTGCTGTGAATGACAGCTTTTGTGAGAAAAATTTTG	1199
QY	661	CCGTGGTTGGACAATTCAGTGAATTTTCATTGAGAAAGAAAGCCTCCATGGAATGTGT	720
Db	1198	CCGTGGTTGGACAATTCAGTGAATTTTCATTGAGAAAGAAAGCCTCCATGGAATGTGT	1257
QY	721	CTAGTGCACGTGTTTGAAGTGGGATCTCCGCGTCCGCGCACATGCGTATCGGCTACATCATG	780
Db	1258	CTAGTGCACGTGTTTGAAGTGGGATCTCCGCGTCCGCGCACATGCGTATGCGCTACATCATG	1317
QY	781	AAAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTACT	840
Db	1318	AAAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTACT	1377
QY	841	ATATCTCCAACTTCAATTTTCTGGGCCAATCTCCCTGGACTATGAGAAAGATTTAAGAAC	900
Db	1378	ATATCTCCAACTTCAATTTTCTGGGCCAATCTCCCTGGACTATGAGAAAGATTTAAGAAC	1437
QY	901	CAGACTGAGATCAGGGCCAAAGAGCAAACTCAAGCGTGTGCACCTGGAAGGCAAT	960
Db	1438	CAGACTGAGATCAGGGCCAAAGAGCAAACTCAAGCGTGTGTGCACCTGGAAGGCAAT	1499
QY	961	GAACCTGTCCCTGCTGTCTCAGAGGGTGTGACAGAAAAAGCGAGCGCCCTCAAGTCAACC	1020
Db	1498	GAACCTGTCCCTGCTGTCTCAGAGGGTGTGACAGAAAAAGCGAGCGCCCTCAAGTCAACC	1557
QY	1021	TGTGCGCACTCTGCTACTCAGAGGCGACAGAACAAAGGCCGTGTGATCTCCGCCAGCGTG	1080
Db	1558	TGTGCGCACTCTGCTACTCAGAGGCGACAGAACAAAGGCCGTGTGATCTCCGCCAGCGTG	1617
QY	1081	CCGACGCGGCCAGCGGTGACGCGGTGTTTAAAGGACACGCGGCTGTATCAAGCGCTC	1148
Db	1618	CCGACGCGGCCAGCGGTGACGCGGTGTTTAAAGGACACGCGGCTGTATCAAGCGCTC	1677
QY	1141	AGTGGGCTGCACCTGTCCGACAGACGGCTGAGACACGCAATTAAGCTCAAGCTTCTCTTC	1200

Db 604 ACTTAGTCCCTACCTGATTCCTCAGCCTTGCTTACCTGTTGGCAACATTGGGCCAAC 663
Qy 481 CGAATTTCTCCCAATCTTATCTTGCGCTGCCAGAGCATGTCTCTCAACAMAGAGCTGATA 540
Db 664 CGAATTTCTCCCAATCTTATCTTGCGCTGCCAGAGCATGTCTCTCAACAMAGAGCTGATA 723
Qy 541 CAGCAGATGGGATTTGGTATGTGTAAATCCAGCTATATACCTGTCCAAAGCCGATCTT 600
Db 724 CAGCAGATGGGATTTGGTATGTGTAAATCCAGCATACTCTTCCAAAGCCGATCTT 783
Qy 601 ATCCCGCATCTCATTTCTGCGGTGCTGTGATGACAGCTTTTGTGAAAAATTTTG 660
Db 784 ATCCCGCATCTCATTTCTGCGGTGCTGTGATGACAGCTTTTGTGAAAAATTTTG 843
Qy 661 CCGTGTGTGACAATCAGTAGATTTCATGGAAGAACAAAGCTTCCATGTGATGTGT 720
Db 844 CCGTGTGTGACAATCAGTAGATTTCATGGAAGAACAAAGCTTCCATGTGATGTGT 903
Qy 721 CTAGTGACATGTTTAGCTGGGATCTCCGCTCCGACCATGCTATCGCTACATCATG 780
Db 904 CTAGTGACATGTTTAGCTGGGATCTCCGCTCCGACCATGCTATCGCTACATCATG 963
Qy 781 AAGGAGATGACATGTCTTTAGATGAAGCTTACAGATTGTGAAAAAGAACCTTACT 840
Db 964 AAGGAGATGACATGTCTTTAGATGAAGCTTACAGATTGTGAAAAAGAACCTTACT 1023
Qy 841 ATATCTCAAACTTCAATTTTCTGGGCCAATCTCTGGACTATGAGAAAGATTAGAAC 900
Db 1024 ATATCTCAAACTTCAATTTTCTGGGCCAATCTCTGGACTATGAGAAAGATTAGAAC 1083
Qy 901 CAGACTGGAGCATGAGGCCAAGAACAACTCAAGCTGTGACCTGGAGAACCAAT 960
Db 1084 CAGACTGGAGCATGAGGCCAAGAACAACTCAAGCTGTGACCTGGAGAACCAAT 1143
Qy 961 GAACTGTCTCTGCTGTCTCAGAGGCTGACAGAAAAAGGACAGCCCTCTCAGTCCACC 1020
Db 1144 GAACTGTCTCTGCTGTCTCAGAGGCTGACAGAAAAAGGACAGCCCTCTCAGTCCACC 1203
Qy 1021 TGTGCGCACTCTGCTCTCAGAGGACAGACAAAGGCCGTGATCCGCGACAGCGTG 1080
Db 1204 TGTGCGCACTCTGCTCTCAGAGGACAGACAAAGGCCGTGATCCGCGACAGCGTG 1263
Qy 1081 CCCAGGCTGCCAGCGCTGACAGCGCTGTGTAGAGACAGCCGCTGTGACAGCGCTC 1140
Db 1264 CCCAGGCTGCCAGCGCTGACAGCGCTGTGTAGAGACAGCCGCTGTGACAGCGCTC 1323
Qy 1141 AGTGGCTGCACTGTCCGACAGAGGCTGGAAGAGAAATAGTCAAGGTTCTTC 1200
Db 1324 AGTGGCTGCACTGTCCGACAGAGGCTGGAAGAGAAATAGTCAAGGTTCTTC 1383
Qy 1201 TCTTGATATCAAAATCAGTTTCAATTCAGCAGCATGGCAGCATCTTACATGCTTC 1260
Db 1384 TCTTGATATCAAAATCAGTTTCAATTCAGCAGCATGGCAGCATCTTACATGCTTC 1443
Qy 1261 TCTTGATCAAGAGCTTTGGAAATCTAACAACTTCCATCTCTGTGATGGACAAAC 1320
Db 1444 TCTTGATCAAGAGCTTTGGAAATCTAACAACTTCCATCTCTGTGATGGACAAAC 1503
Qy 1321 AAGCTATGCGAGTTCTCCCTGTTCAGAGAACTATGGAGACATCCCGAAACCATGCTCT 1380
Db 1504 AAGCTATGCGAGTTCTCCCTGTTCAGAGAACTATGGAGACATCCCGAAACCATGCTCT 1563
Qy 1381 GATTAAAGAGAAACAGCATCCCAAGAAAGCTGACAGCCGACGCTTTCAGACAGCCAG 1440
Db 1564 GATTAAAGAGAAACAGCATCCCAAGAAAGCTGACAGCCGACGCTTTCAGACAGCCAG 1623
Qy 1441 AGCAGAGCATTTGGTGTGAGAACAGCAGCATGTGGACCGCCAGAGGTCCCTTTTA 1500
Db 1624 AGCAGAGCATTTGGTGTGAGAACAGCAGCATGTGGACCGCCAGAGGTCCCTTTTA 1683
Qy 1501 TCTTCATGCAATGCAAGTGGAGCTGGAGGACATTTACACACAGCTTCTTTTGGGC 1560
Db 1684 TCTTCATGCAATGCAAGTGGAGCTGGAGGACATTTACACACAGCTTCTTTTGGGC 1743

Qy 1561 CTTTCACAGCCAGACAGCCTTCAAGATGCTGTGCGCTGGACCTTAAAGGCTGGAC 1620
Db 1744 CTTTCACAGCCAGACAGCCTTCAAGATGCTGTGCGCTGGACCTTAAAGGCTGGAC 1803
Qy 1621 TGGGATATCTTGGCCCCCAGACCTCTACCCCTTCTGACCGACAGCTGTATTTTGGC 1680
Db 1804 TGGGATATCTTGGCCCCCAGACCTCTACCCCTTCTGACCGACAGCTGTATTTTGGC 1863
Qy 1681 ACAGAGTCTCACAATCTCTACCTGCTCAGCATCTACGAGGACGAGGACGCTTACTCT 1740
Db 1864 ACAGAGTCTCACAATCTCTACCTGCTCAGCATCTACGAGGACGAGGACGCTTACTCT 1923
Qy 1741 GCTTACAGTGCAGCCAGCTGCCCATTTTGGGAGACCAAGCTATTTCTGTGCGAGCGG 1800
Db 1924 GCTTACAGTGCAGCCAGCTGCCCATTTTGGGAGACCAAGCTATTTCTGTGCGAGCGG 1983
Qy 1801 CAGAGCCAGAGTGAAGAGCTGACTCGGGGGAGAGCTGGCATGAAAGAGCCCTTTGAA 1860
Db 1984 CAGAGCCAGAGTGAAGAGCTGACTCGGGGGAGAGCTGGCATGAAAGAGCCCTTTGAA 2043
Qy 1981 AAGCATTTAAACGACAGAGCTGCCAAATGGAATTTGAGAGACATCATGTACAGAAC 1920
Db 2044 AAGCATTTAAACGACAGAGCTGCCAAATGGAATTTGAGAGACATCATGTACAGAAC 2103
Qy 1921 AGTTCACGGGAAGAGCTGGGAAAGTGGGACGTAGCTATTTTGGGACAGCATGAA 1980
Db 2104 AGTTCACGGGAAGAGCTGGGAAAGTGGGACGTAGCTATTTTGGGACAGCATGAA 2163
Qy 1981 ATCATTTAGGTCTCC 1995
Db 2164 ATCATTTAGGTCTCC 2178

RESULT 9
ABV20833
ID ABV20833 standard; cDNA; 5145 BP.
XX
AC ABV20833;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 20824.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN W02001.60860-A2.
XX
PD 23-AUG-2001.
XX
PE 20-FEB-2001; 2001MO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3419; 11750bp; English.

XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 94.9%; Score 1893; DB 5; Length 5145;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGTATGTTGAATCAATTTGTTACTGAGAGTGTGGTCTGCTGGA 60
DB ATGGCCCATGAGTATGTTGAATCAATTTGTTACTGAGAGTGTGGTCTGCTGGA 648
QY 61 AGTGAACGGAAAAAGTGTCTAATTGATGAGCGGCAATTGTGGAAATACATACATCC 120
DB AGTGAACGGAAAAAGTGTCTAATTGATGAGCGGCAATTGTGGAAATACATACATCC 708
QY 121 CACATTTTGGAAACCATTAATATCACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 180
DB CACATTTTGGAAACCATTAATATCACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 768
QY 181 GACAAAGTGTATTAACAGAGCTCATTCAGCATTCAGCGAAATATAGTTGACATTAT 240
DB GACAAAGTGTATTAACAGAGCTCATTCAGCATTCAGCGAAATATAGTTGACATTAT 828
QY 241 TGCAGTCAGAAAGTTGATGATTCAGATCAAAAGCTCCCAAGATGTGCTCTCTCTCA 300
DB TGCAGTCAGAAAGTTGATGATTCAGATCAAAAGCTCCCAAGATGTGCTCTCTCTCA 888
QY 829 TGCAGTCAGAAAGTTGATGATTCAGATCAAAAGCTCCCAAGATGTGCTCTCTCTCA 888
QY 301 GACTGTTTCTCACTGACTTCTGGGTAACCTGAGAAAGCTTCAACTCTGTTCACTG 360
DB GACTGTTTCTCACTGACTTCTGGGTAACCTGAGAAAGCTTCAACTCTGTTCACTG 948
QY 361 CTTCGAGTGGGTTTGTGTAAGTCTCTGTTGTTTCCCTGCTCTGTGAAGAAATCC 420
DB CTTCGAGTGGGTTTGTGTAAGTCTCTGTTGTTTCCCTGCTCTGTGAAGAAATCC 1008
QY 421 ACTCTAGTCCCTTACCTGCAATTTCTCAGCCTTGCTTACCTGTTGCCAAACATTGGGCCAAC 480
DB ACTCTAGTCCCTTACCTGCAATTTCTCAGCCTTGCTTACCTGTTGCCAAACATTGGGCCAAC 1068
QY 1009 ACTCTAGTCCCTTACCTGCAATTTCTCAGCCTTGCTTACCTGTTGCCAAACATTGGGCCAAC 1068
QY 481 CGAATTTCTTCCCATCTTTATCTTGGCTGCGCAGAGATGTCTTCAACAGAGCTGATA 540
DB CGAATTTCTTCCCATCTTTATCTTGGCTGCGCAGAGATGTCTTCAACAGAGCTGATA 1128
QY 1069 CGAATTTCTTCCCATCTTTATCTTGGCTGCGCAGAGATGTCTTCAACAGAGCTGATA 1128
QY 541 CAGCAGAAATGAGTGGTATGTTGTTAATGCGACCTATACCTGTGCCAAAGCTGACTT 600
DB CAGCAGAAATGAGTGGTATGTTGTTAATGCGACCTATACCTGTGCCAAAGCTGACTT 1188
QY 1129 CAGCAGAAATGAGTGGTATGTTGTTAATGCGACCTATACCTGTGCCAAAGCTGACTT 1188
QY 601 ATCCCGAGTCTCATTTCTGCGCTGCGCTGCGTGAATGACAGCTTTGTGGAATAATTTTG 660
DB ATCCCGAGTCTCATTTCTGCGCTGCGCTGCGTGAATGACAGCTTTGTGGAATAATTTTG 1248
QY 1189 ATCCCGAGTCTCATTTCTGCGCTGCGCTGCGTGAATGACAGCTTTGTGGAATAATTTTG 1248
QY 661 CCGTGGTTGGAACAATCAGATGATTTCAATGAGAAAGCAAAAGCTTCAATGATGTT 720
DB CCGTGGTTGGAACAATCAGATGATTTCAATGAGAAAGCAAAAGCTTCAATGATGATGTT 1308
QY 1249 CCGTGGTTGGAACAATCAGATGATTTCAATGAGAAAGCAAAAGCTTCAATGATGATGTT 1308
QY 721 CTAGTGCATGTTTATGTTGAGTGTCTCCGCTCCGCCACCATGCTATGCTTACATCATG 780
DB CTAGTGCATGTTTATGTTGAGTGTCTCCGCTCCGCCACCATGCTATGCTTACATCATG 1368
QY 1309 CTAGTGCATGTTTATGTTGAGTGTCTCCGCTCCGCCACCATGCTATGCTTACATCATG 1368
QY 781 AAGAGATGAGCATGCTTATGATGAGCTTACAGATTTGTGAAGAAAAAAGCACTACT 840

DB 1369 AAGAGATGAGCATGCTTATGATGAGCTTACAGATTTGTGAAGAAAAAAGCACTACT 1428
QY 841 ATATCTCCAAATTTTCTTGGGCGCACTCTCTGAGATATGAGAAAGATTAGAAC 900
DB 1429 ATATCTCCAAATTTTCTTGGGCGCACTCTCTGAGATATGAGAAAGATTAGAAC 1488
QY 901 CAGACTGAGCATCAGGGCGCAAGAGCAAACTCAAGCTCTGACCTGAGAAAGCCAAAT 960
DB 1489 CAGACTGAGCATCAGGGCGCAAGAGCAAACTCAAGCTCTGACCTGAGAAAGCCAAAT 1548
QY 961 GAACCTGTCTCTGCTGTCTCAGAGGGTGGACAGAAAGGAGAGCGCCCTCAGTCCACC 1020
DB 1549 GAACCTGTCTCTGCTGTCTCAGAGGGTGGACAGAAAGGAGAGCGCCCTCAGTCCACC 1608
QY 1021 TGTGCCGACTGTCTCTCAGAGGGTGGACAGAAAGGAGAGCGCCCTCAGTCCACC 1080
DB 1609 TGTGCCGACTGTCTCTCAGAGGGTGGACAGAAAGGAGAGCGCCCTCAGTCCACC 1668
QY 1081 CCCAGGTCGCCAGGGTGGACGCTGCTGTTAGAGAGCAGGCCGCTGTACAGGGCTTC 1140
DB 1669 CCCAGGTCGCCAGGGTGGACGCTGCTGTTAGAGAGCAGGCCGCTGTGTACAGGGCTTC 1728
QY 1141 AGTGGCTGCACTGTCTGCGCAGACAGGCTGGAAGACAGCAATAGCTCAAGGCTTCTTC 1200
DB 1729 AGTGGCTGCACTGTCTGCGCAGACAGGCTGGAAGACAGCAATAGCTCAAGGCTTCTTC 1788
QY 1201 TCTCTGATATCAATCAATCAATTCATTCAGCCAGATGAGCATCTTACATGCTTC 1260
DB 1789 TCTCTGATATCAATCAATCAATTCATTCAGCCAGATGAGCATCTTACATGCTTC 1848
QY 1261 TCTCTGATGAGATGCTTGGAAATCTCAAACTTCCACTGATGATGGAGCAAC 1320
DB 1849 TCTCTGATGAGATGCTTGGAAATCTCAAACTTCCACTGATGATGGAGCAAC 1908
QY 1321 AAGCTATGCAATTTCTCCCTGTTGAGAACTATCGAGACAGATCTCCGAAACCACTCT 1380
DB 1909 AAGCTATGCAATTTCTCCCTGTTGAGAACTATCGAGACAGATCTCCGAAACCACTCT 1968
QY 1381 GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGACAGCCGACAGCTTACAGACCCAG 1440
DB 1969 GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGACAGCCGACAGCTTACAGACCCAG 2028
QY 1441 AGCAAGCGATTCATTCCTGCTGAGAACCAACAGAGTGGACCGCCAGAGGTCCTTTTA 1500
DB 2029 AGCAAGCGATTCATTCCTGCTGAGAACCAACAGAGTGGACCGCCAGAGGTCCTTTTA 2088
QY 1501 TCTCCACTGATGAGAGTGGAGCGTGGAGCAATTTACACACAGCTTCTTTTGGG 1560
DB 2089 TCTCCACTGATGAGAGTGGAGCGTGGAGCAATTTACACACAGCTTCTTTTGGG 2148
QY 1561 CTTTCCACAGCAGCAGCAGCAGCTTCAAGAGTGTGCTGGCTTAAAGGCTGGCAC 1620
DB 2149 CTTTCCACAGCAGCAGCAGCAGCTTCAAGAGTGTGCTGGCTTAAAGGCTGGCAC 2208
QY 1621 TGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGC 1680
DB 2209 TGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGC 2268
QY 1681 ACAGAGTCTCACTTACTCTGCTCAGCAGCATCTACGAGGAGGAGTCCAGTACTCT 1740
DB 2269 ACAGAGTCTCACTTACTCTGCTCAGCAGCATCTACGAGGAGGAGTCCAGTACTCT 2328
QY 1741 GCTTACAGTGCAGCCAGCTGCCCTTCCAGAGCAAGTATTTCTGTGCGCAGGCG 1800
DB 2329 GCTTACAGTGCAGCCAGCTGCCCTTCCAGAGCAAGTATTTCTGTGCGCAGGCG 2388
QY 1801 CAGAGCCCAAGTGCAGAGCTGATCTGCGCGGAGAGTGGCATGAAAGAGCCCTTTGAA 1860
DB 2389 CAGAGCCCAAGTGCAGAGCTGATCTGCGCGGAGAGTGGCATGAAAGAGCCCTTTGAA 2448
QY 1861 AAGCAGTTTAAAGCAGAGAGTGCCTCAATGAAATTTGAGAGAGCATGCTCAGAGAAC 1920

Db 2449 AAGCAGTTAAACGAGAGCTGCCAAATGGAATTTGGAGAGCATCATGTACAGAAC 2508
QY 1921 AGTCACGGGAAAGAGCTGGGAAAGATGGGCACTCAGTTACTTTTGGGCGACATGAA 1980
Db 2509 AGGTACGGGAAAGAGCTGGGAAAGATGGGCACTCAGTTACTTTTGGGCGACATGAA 2568
QY 1981 ATCATTTAGGTCTCC 1995
Db 2569 ATCATTTAGGTCTCC 2583

RESULT 10
ABV21080
ID ABV21080 standard; cDNA; 5145 BP.
AC ABV21080;
XX
XX 13-SEP-2002 (first entry)
DT
XX Human prostate expression marker cDNA 2107L.
DE
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
OS Homo sapiens.
XX
XX WO200160860-A2.
PN
XX 23-AUG-2001.
PD
XX 20-FEB-2001; 2001WO-US005171.
PF
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-025281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
PI Schlegel R, Endege WO, Monahan UE;
PI WPI; 2001-662795/76.
DR
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 3481; 11750pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC of progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 94.9%; Score 1893; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGGAATCAATTTGTACTGAGAGTTGGTGGCTCTGGTAAA 60
Db 589 ATGGCCCATGAGATGATGGAATCAATTTGTACTGAGAGTTGGTGGCTCTGGTAAA 648

QY 61 AGTGAACCGAAAAAGTGTGCTAAATGATAGCCGGCACTTTGTGAATATCAATCATCC 120
Db 649 AGTGAACCGAAAAAGTGTGCTAAATGATAGCCGGCACTTTGTGAATATCAATCATCC 708
QY 121 CACATTTTGGAAAGCCATTATATCACTGTCTCAAGCTTTATGAAAGCGAAGTTGCAAG 180
Db 709 CACATTTTGGAAAGCCATTATATCACTGTCTCAAGCTTTATGAAAGCGAAGTTGCAAG 768
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCAATTCAGGAAACATTAAGTTGACATTAAT 240
Db 769 GACAAAGTGTAAATTAACAGAGCTCATCCAGCAATTCAGGAAACATTAAGTTGACATTAAT 828
QY 241 TGCAGTCAAGAAAGTTGATTAACATCAAGAGCTCCCAAGATGTTGCTCTCTCTCA 300
Db 829 TGCAGTCAAGAAAGTTGATTAACATCAAGAGCTCCCAAGATGTTGCTCTCTCTCA 888
QY 301 GACTGTTTTTCACTGTACTTCTGGGTAACTGGAAGAGCTTCAACTCTGTTACCTG 360
Db 889 GACTGTTTTTCACTGTACTTCTGGGTAACTGGAAGAGCTTCAACTCTGTTACCTG 948
QY 361 CTGACAGGTGGGTTTGTGAGATTCTCTGTTGTTCCCTGGGCTCTGTGAAGAAATATCC 420
Db 949 CTGACAGGTGGGTTTGTGAGATTCTCTGTTGTTCCCTGGGCTCTGTGAAGAAATATCC 1008
QY 421 ACTGTAGTCCCTACCTGATTTCTCAGCCTTGTCTTACCTGTGCAACATTTGGGCAACC 480
Db 1009 ACTGTAGTCCCTACCTGATTTCTCAGCCTTGTCTTACCTGTGCAACATTTGGGCAACC 1068
QY 481 CGAATTTCTCCCAATCTTTATCTTGGGCGCAGGAAATATCTCTCAACAGAGCTATA 540
Db 1069 CGAATTTCTCCCAATCTTTATCTTGGGCGCAGGAAATATCTCTCAACAGAGCTATA 1128
QY 541 CAGCAGAAATGGATTTGTTATGTTTAAATGCCAGTATACCTGTCAAGAGCTGACTTT 600
Db 1129 CAGCAGAAATGGATTTGTTATGTTTAAATGCCAGTATACCTGTCAAGAGCTGACTTT 1188
QY 601 ATCCCCGAGTCTCAATTTCTGCGTGTGCTGTGATGACAGCTTTGTGAAAAATTTTG 660
Db 1189 ATCCCCGAGTCTCAATTTCTGCGTGTGCTGTGATGACAGCTTTGTGAAAAATTTTG 1248
QY 661 CCGTGTGTTGACAATCAGTAGATTTTCATTTGAGAAAGCAAAAGCCCTCAATGGAATGTT 720
Db 1249 CCGTGTGTTGACAATCAGTAGATTTTCATTTGAGAAAGCAAAAGCCCTCAATGGAATGTT 1308
QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCAACATCTGTTATGCTTACATCATG 780
Db 1309 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCAACATCTGTTATGCTTACATCATG 1368
QY 781 AAGAGATGAGCATGCTTTTAAATGAAAGCTTACAGATTTGTGAAAGAAAAAGACTTACT 840
Db 1369 AAGAGATGAGCATGCTTTTAAATGAAAGCTTACAGATTTGTGAAAGAAAAAGACTTACT 1428
QY 841 ATATCTCAAAATCTCAATTTTCTGGGCAATCTCTGCACTATGAGAAAGATTTAAGAAC 900
Db 1429 ATATCTCAAAATCTCAATTTTCTGGGCAATCTCTGCACTATGAGAAAGATTTAAGAAC 1488
QY 901 CAACTGAGAGCATCAGGCGCAAAAGCAAACTTAAGCTCTGACCTGGAAGAGCAAT 960
Db 1489 CAACTGAGAGCATCAGGCGCAAAAGCAAACTTAAGCTCTGACCTGGAAGAGCAAT 1548
QY 961 GAACCTGTCCCTGCTGTCAAGAGGTGAGCAAGAAAGGAAAGCCCTCTGAGTCAACC 1020
Db 1549 GAACCTGTCCCTGCTGTCTCAAGAGGTGAGCAAGAAAGGAAAGCCCTCTGAGTCAACC 1608
QY 1021 TGTGCGACTTGTCTAATCTCAGAGCAGAGCAAGAGCCGCTGATCCGCGACGCTG 1080
Db 1609 TGTGCGACTTGTCTAATCTCAGAGCAGAGCAAGAGCCGCTGATCCGCGACGCTG 1668
QY 1081 CCCAGGTGCTCCAGGCTGAGCCGCTGCTGTTAAGAGAACAGCCGCTGTTACAGGCGCTC 1140
Db 1669 CCCAGGTGCTCCAGGCTGAGCCGCTGCTGTTAAGAGAACAGCCGCTGTTACAGGCGCTC 1728

Db 949 CTTCAGAGTGGGTTTGCTGAGTTCTCTGTTGTTTCCCTGGGCTCTGTGAGGAATAATCC 1008
Qy 421 ACTTAGTCCCTACCTGATTTCTCAGGCTTGCTGCTACTGTTGGCAATTTGGGCCAAC 480
Db 1009 ACTTAGTCCCTACCTGATTTCTCAGGCTTGCTGCTACTGTTGGCAATTTGGGCCAAC 1068
Qy 481 CGAATTTCTCCCAATCTTTATCTTTGCTGCGCAGGAGATGTCTCAACAAGAGCTGTATA 540
Db 1069 CGAATTTCTCCCAATCTTTATCTTTGCTGCGCAGGAGATGTCTCAACAAGAGCTGTATA 1128
Qy 541 CAGAGAGATGGAGTGGTTATGTGTATAATGCCAGCTATACCTGTCCAAAGCTGACTT 600
Db 1129 CAGAGAGATGGAGTGGTTATGTGTATAATGCCAGCTATACCTGTCCAAAGCTGACTT 1188
Qy 601 ATCCCCAGTCTCATTTCTGCGTGTGCTGTGATGACAGCTTTTGTGAGAAATTTTG 660
Db 1189 ATCCCCAGTCTCATTTCTGCGTGTGCTGTGATGACAGCTTTTGTGAGAAATTTTG 1248
Qy 661 CCGTGGTTGGCAAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTGT 720
Db 1249 CCGTGGTTGGCAAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTGT 1308
Qy 721 CTAGTGACCTGTTAGTGGGATCTCCGCTCCGCCACCATGCTTATCCGCTATCATCATG 780
Db 1309 CTAGTGACCTGTTAGTGGGATCTCCGCTCCGCCACCATGCTTATCCGCTATCATCATG 1368
Qy 781 AAGAGATGACATGCTCTTATAGTAGAGCTTACAGATTTGTGAAAAGAAAAGCACTACT 840
Db 1369 AAGAGATGACATGCTCTTATAGTAGAGCTTACAGATTTGTGAAAAGAAAAGCACTACT 1428
Qy 841 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTGACCTATGAGAAAGATTAGAAC 900
Db 1429 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTGACCTATGAGAAAGATTAGAAC 1488
Qy 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAACTGCTGACCTGTGAGAAAGCCAAAT 960
Db 1489 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAACTGCTGACCTGTGAGAAAGCCAAAT 1548
Qy 961 GAACTGTCTCTGCTGTCTCAGAGGCTGAGCAGAAAAGGAGAGAGCCCTCAGTCCACCC 1020
Db 1549 GAACTGTCTCTGCTGTCTCAGAGGCTGAGCAGAAAAGGAGAGAGCCCTCAGTCCACCC 1608
Qy 1021 TGTGCGGACTTGTCTTACTCAGAGGCGACAGACAAAGGCCGTGATCCGCCAGCGTG 1080
Db 1609 TGTGCGGACTTGTCTTACTCAGAGGCGACAGACAAAGGCCGTGATCCGCCAGCGTG 1668
Qy 1081 CCCAGGCTGCCAGGCTGTGAGCGGTGCTTAAAGGAGAGAGCCCGCTGTATAGGCGCTC 1140
Db 1669 CCCAGGCTGCCAGGCTGTGAGCGGTGCTTAAAGGAGAGAGCCCGCTGTATAGGCGCTC 1728
Qy 1141 AGTGGGCTGCACCTGTCCGACAGAGGCTGAAAGCAGCAATTAAGCTCAAGGCTTCTTC 1200
Db 1729 AGTGGGCTGCACCTGTCCGACAGAGGCTGAAAGCAGCAATTAAGCTCAAGGCTTCTTC 1788
Qy 1201 TCTCTGAGATCAAAATCAGTTTCAATTTAGCCAGCATGCGAGCATCTTTCATGAGCTTC 1260
Db 1789 TCTCTGAGATCAAAATCAGTTTCAATTTAGCCAGCATGCGAGCATCTTTCATGAGCTTC 1848
Qy 1261 TCCCTATCAGAGATGCTTTGGATATCTAACAACCTTCCACTCATCTGAGAGGAGCAAC 1320
Db 1849 TCCCTATCAGAGATGCTTTGGATATCTAACAACCTTCCACTCATCTGAGAGGAGCAAC 1908
Qy 1321 AAGCTATGCGAGTTCTCCCTGTGTCAGGAACTATCGAGAGCATCTCCGAAACCATGCTCT 1380
Db 1909 AAGCTATGCGAGTTCTCCCTGTGTCAGGAACTATCGAGAGCATCTCCGAAACCATGCTCT 1968
Qy 1381 GATTAGAGGAAAGCCAGCATCTCCCAAGAACTGCGACCGCCAGGCTTTCAGACGCCAG 1440
Db 1969 GATTAGAGGAAAGCCAGCATCTCCCAAGAACTGCGACCGCCAGGCTTTCAGACGCCAG 2028
Qy 1441 AGCAAGCATTTGCTTGGTGTGAGACCGAGCAGCAGTGGACCGGCCAAGAGTCCCTTTTA 1500
Db 2029 AGCAAGCATTTGCTTGGTGTGAGACCGAGCAGCAGTGGACCGGCCAAGAGTCCCTTTTA 2088

Qy 1501 TCTCCACTGATCCGAAGTGGAGCGGTGAGAGCAATTAACACACCAAGCTTCTTTCCGC 1560
Db 2089 TCTCCACTGATCCGAAGTGGAGCGGTGAGAGCAATTAACACACCAAGCTTCTTTCCGC 2148
Qy 1561 CTTTCCACGAGCAGACAGCACTTCAAGAGTGTGTGCTGGGCTTAAAGGCTGGGAC 1620
Db 2149 CTTTCCACGAGCAGACAGCACTTCAAGAGTGTGTGCTGGGCTTAAAGGCTGGGAC 2208
Qy 1621 TCGATATCTTGGCCCCCAAGCTTACCCCTTCCCTGACACGACGCTGTATTTTGC 1680
Db 2209 TCGATATCTTGGCCCCCAAGCTTACCCCTTCCCTGACACGACGCTGTATTTTGC 2268
Qy 1681 ACAGAGTCTTCACTTACTCTGCTGCTGACGACATCTAAGAGAGAGTCCAGTACTCT 1740
Db 2269 ACAGAGTCTTCACTTACTCTGCTGCTGACGACATCTAAGAGAGAGTCCAGTACTCT 2328
Qy 1741 GCTTACAGCTGACGACGACCTGCCACTTGGCGAGACCAAGCTTATCTGTGCGAGGCG 1800
Db 2329 GCTTACAGCTGACGACGACCTGCCACTTGGCGAGACCAAGCTTATCTGTGCGAGGCG 2388
Qy 1801 CAGAGCCAAAGTACAGAGCTGACTTGGCGGAGGAGCTGGCATGAAGAGAGCCCTTTGAA 1860
Db 2389 CAGAGCCAAAGTACAGAGCTGACTTGGCGGAGGAGCTGGCATGAAGAGAGCCCTTTGAA 2448
Qy 1861 AAGCATTTAAACGAGAAAGCTGCAAAATGGAATTTGAGAGACATCATGTGAGAAC 1920
Db 2449 AAGCATTTAAACGAGAAAGCTGCAAAATGGAATTTGAGAGACATCATGTGAGAAC 2508
Qy 1921 AGGTACCGGAAAGAGCTGGGAAAGTGGGAGTCAAGTCTTATGCGGAGCATGGA 1980
Db 2509 AGGTACCGGAAAGAGCTGGGAAAGTGGGAGTCAAGTCTTATGCGGAGCATGGA 2568
Qy 1981 ATCATTTAGGCTCTCC 1995
Db 2569 ATCATTTAGGCTCTCC 2583

RESULT 12
ABV20978
ID ABV20978 standard; cDNA; 5145 BP.
XX
XX ABV20978;
AC
XX
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 20969.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200160860-A2.
PD
PD 23-AUG-2001.
PE
PE 20-FEB-2001; 2001MO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 16-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE.
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1, Page 3451, 11750pp, English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC of progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC / (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 94.98; Score 1893; DB 5; Length 5145;

Best Local Similarity 99.98; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGCCCATGAGATGATGAGACTCAATGTTACTGAGAGGTGGTGGCTCTGCGGAA 60
DB 589 ATGACCATGAGATGATGAGACTCAATGTTACTGAGAGGTGGTGGCTCTGCGGAA 648
QY 61 AGTGAACGGAAGAGTGTCTAATGATAGCGGCGCATTTGTGAAATACATATACCTC 120
DB 649 AGTGAACGGAAGAGTGTCTAATGATAGCGGCGCATTTGTGAAATACATATACCTC 708
QY 121 CACATTTGGAAGCATTAAATATCACTGCTCCAGTTATGAAGCGAAGTTGCAACAG 180
DB 709 CACATTTGGAAGCATTAAATATCACTGCTCCAGTTATGAAGCGAAGTTGCAACAG 768
QY 181 GACAAAGTGTATTAAGAGCTCATACGACATTCAGCGAAACATTAAGTTGACATTGAT 240
DB 769 GACAAAGTGTATTAAGAGCTCATACGACATTCAGCGAAACATTAAGTTGACATTGAT 828
QY 241 TGCAGTCAGAGGTGTAGTTTACGATCAAGCTCCCAAGATGTGCTCTCTCTTCA 300
DB 829 TGCAGTCAGAGGTGTAGTTTACGATCAAGCTCCCAAGATGTGCTCTCTCTTCA 888
QY 301 GACTGTTTCTCACTGACTTCTGGGTAACCTGAGAAAGCTTCAACTCTGTTACCTG 360
DB 889 GACTGTTTCTCACTGACTTCTGGGTAACCTGAGAAAGCTTCAACTCTGTTACCTG 948
QY 361 CTTCAGAGTGGGTTTGTGAGTTCTCTCGTTGTTTCCCTGGCTCTGTGAAAGAAATCC 420
DB 949 CTTCAGAGTGGGTTTGTGAGTTCTCTCGTTGTTTCCCTGGCTCTGTGAAAGAAATCC 1008
QY 421 ACTCTAGTCCCTACCTGATCTTCTCAGCTTGTGCTTACCTGTGCAACATTTGGGCCAAC 480
DB 1009 ACTCTAGTCCCTACCTGATCTTCTCAGCTTGTGCTTACCTGTGCAACATTTGGGCCAAC 1068
QY 481 CGAATTTCTTCCCAATCTTAACTTGTGCTGCGCAGCGAGATGTCTTCAACAGAGCTGATA 540
DB 1069 CGAATTTCTTCCCAATCTTAACTTGTGCTGCGCAGCGAGATGTCTTCAACAGAGCTGATA 1128
QY 541 CACCAAGATGGATGTGTTATGTGTTAAATGCCAGCTATACCTGTCCAAAGGCTGACTTT 600
DB 1129 CACCAAGATGGATGTGTTATGTGTTAAATGCCAGCTATACCTGTCCAAAGGCTGACTTT 1188
QY 601 ATCCCGAGTCTCATTTCTGCGGTGCTGCTGATGACAGCTTTTGTGAAATTTTGG 660
DB 1189 ATCCCGAGTCTCATTTCTGCGGTGCTGCTGATGACAGCTTTTGTGAAATTTTGG 1248
QY 661 CCGTGTGTGACAAATCAGTAGATTTCATTGAGAAACAAAGCTTCAATGATGTGT 720
DB 1249 CCGTGTGTGACAAATCAGTAGATTTCATTGAGAAACAAAGCTTCAATGATGTGT 1308
QY 721 CTAGTGCATGTTAGTGGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATG 780
```

```
DB 1309 CTAGTGCATGTTAGTGGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATG 1368
QY 781 AAGAGATGAGACATGTCTTATAGTAGACTTACAGATTTGTGAAAGAAAGAAAGCACTACT 840
DB 1369 AAGAGATGAGACATGTCTTATAGTAGACTTACAGATTTGTGAAAGAAAGAAAGCACTACT 1428
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCGCAACTCTGAGCTATAGAAAGAAATTAAGAAC 900
DB 1429 ATATCTCCAAACTTCAATTTTCTGGGCGCAACTCTGAGCTATAGAAAGAAATTAAGAAC 1488
QY 901 CAGACTGAGACATCAGGCGCAAGAGCAAACTCAAGCTGTGACCTGAGAAAGCCAAAT 960
DB 1489 CAGACTGAGACATCAGGCGCAAGAGCAAACTCAAGCTGTGACCTGAGAAAGCCAAAT 1548
QY 961 GAACTGTCTCTGCTGTCTTCAAGAGGTGACAGAAAGCGAGACGCCCTCACTCAACC 1020
DB 1549 GAACTGTCTCTGCTGTCTTCAAGAGGTGACAGAAAGCGAGACGCCCTCACTCAACC 1608
QY 1021 TGTGCGGACTGTGCTTACCTCAGAGGACAGACAAAGGCGCTGCATCCCGCAGCGTG 1080
DB 1609 TGTGCGGACTGTGCTTACCTCAGAGGACAGACAAAGGCGCTGCATCCCGCAGCGTG 1668
QY 1081 CCCAGGCGTCCAGCGTGCAGCGGCTGCTGTAGAGAGACAGCGCGCTGATACAGGCGCTC 1140
DB 1669 CCCAGGCGTCCAGCGTGCAGCGGCTGCTGTAGAGAGACAGCGCGCTGATACAGGCGCTC 1728
QY 1141 AGTGGCTGCACCTGTCCGACAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
DB 1729 AGTGGCTGCACCTGTCCGACAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1788
QY 1201 TCTCTGATATCAATCAGTTTATATTCAGGACAGATGACAGATCTTACATGTGCTTC 1260
DB 1789 TCTCTGATATCAATCAGTTTATATTCAGGACAGATGACAGATCTTACATGTGCTTC 1848
QY 1261 TCCTCATCAGAAAGATCTTTGGAATCTAACAACCTTCCACTCTGTGATGGAAGCAAC 1320
DB 1849 TCCTCATCAGAAAGATCTTTGGAATCTAACAACCTTCCACTCTGTGATGGAAGCAAC 1398
QY 1321 AAGCTATGCGCAGTTCTCCCTGTTCAGGAATATCGAGAGCAGCTCCGAAACAGTCTCT 1380
DB 1909 AAGCTATGCGCAGTTCTCCCTGTTCAGGAATATCGAGAGCAGCTCCGAAACAGTCTCT 1968
QY 1381 GATTAAGAGAGAGCCAGCATCTCCCAAGAACTGCAAGCGGCTTACAGACAGCAG 1440
DB 1969 GATTAAGAGAGAGCCAGCATCTCCCAAGAACTGCAAGCGGCTTACAGACAGCAG 2028
QY 1441 AGCAAGCATTTGATTCGGTCAGAAACAGAGCAGGAGCGCCAGAGGTCCTTTTAA 1500
DB 2029 AGCAAGCATTTGATTCGGTCAGAAACAGAGCAGGAGCGCCAGAGGTCCTTTTAA 2088
QY 1501 TCTCAGCTGATGAGAGTGGAGACGTGAGAGCAATTAACACACAGCTTCTTTTGGCG 1560
DB 2089 TCTCAGCTGATGAGAGTGGAGACGTGAGAGCAATTAACACACAGCTTCTTTTGGCG 2148
QY 1561 CTTTCAACAGAGCAGAGCAGCATCTCAGAAAGTCTGTGCGCTGTGAGGCTGTGAC 1620
DB 2149 CTTTCAACAGAGCAGAGCAGCATCTCAGAAAGTCTGTGCGCTGTGAGGCTGTGAC 2208
QY 1621 TCGGATATCTTGGCCCCCGACAGCTTACCCCTTCCCTGACAGAGCTGTATTTTGGC 1680
DB 2209 TCGGATATCTTGGCCCCCGACAGCTTACCCCTTCCCTGACAGAGCTGTATTTTGGC 2268
QY 1681 ACAGAGTCTCTCACTTCTACTGTGCTGAGCCATAGAGAGAGAGTGCACATTACTCT 1740
DB 2269 ACAGAGTCTCTCACTTCTACTGTGCTGAGCCATTAAGAGAGAGTGCACATTACTCT 2328
QY 1741 GCTTAAGCTGACAGCAGCTGCTGCCATTTGGGAGACCAAGTCTATTCTGTGCGCAGCGG 1800
DB 2329 GCTTAAGCTGACAGCAGCTGCTGCCATTTGGGAGACCAAGTCTATTCTGTGCGCAGCGG 2388
QY 1801 CAGAGCAGAGTGAAGAGCTGACTGCGGCGGAGCTGCGATGAAGAGGCCCTTTGAA 1860
```

Db 2389 CAGAGCCAAAGTACAGAGCTGACTCCGCGGAGAGTGCATGAGAGAGCCCTTTGAA 2448
QY 1861 AAGCAGTTTAAAGCAGAGCTGCCAATGGAAATTGGAGAGACATCATGTAGAGAAC 1920
Db 2449 AAGCAGTTTAAAGCAGAGCTGCCAATGGAAATTGGAGAGACATCATGTAGAGAAC 2508
QY 1921 AGGTACGCGGAGAGAGTGGGAAAAGTGGGAGTCACTCTTTTGGGAGACATGAA 1980
Db 2509 AGGTACGCGGAGAGAGTGGGAAAAGTGGGAGTCACTCTTTTGGGAGACATGAA 2568
QY 1981 ATCATTTAGAGTCTCC 1995
Db 2569 ATCATTTAGAGTCTCC 2583

RESULT 13
ABV21092
ID ABV21092 standard; cDNA; 5145 BP.
AC ABV21092;
DT 13-SEP-2002 (first entry)
DE Human prostate expression marker cDNA 21083.
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
OS Homo sapiens.
PN MO200160860-A2.
PD 23-AUG-2001.
XX 20-FEB-2001; 2001MO-US005171.
XX 17-FEB-2000; 2000US-0183119P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIDM PREDICTIVE MEDICINE INC.
XX PA Schlegel R, Endege WO, Monahan JE;
XX PI WPI; 2001-662795/76.
XX DR
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1, Page 3485; 11750pp; English.
XX
XX CC The invention relates to an isolated nucleic acid molecule (1) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
XX CC specification or its complement. (1) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 94.9%; Score 1893; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATTGGAATCAAAATTGTTACTGAGAGGTGGTCTCTGTCGAA 60
Db 589 ATGGCCCATGAGATGATTGGAATCAAAATTGTTACTGAGAGGTGGTCTCTGTCGAA 648
QY 61 AGTGGAAACGGAAGAAAGTGTCTGTAATGTAAGCCGACATTTGGTAATATCAATCATCC 120
Db 649 AGTGGAAACGGAAGAAAGTGTCTGTAATGTAAGCCGACATTTGGTAATATCAATCATCC 708
QY 121 CACATTTTGGAAAGCCATTATATATCAATGCTTCCAAAGCTTATGAAAGGTTGCAACG 180
Db 709 CACATTTTGGAAAGCCATTATATCAATGCTTCCAAAGCTTATGAAAGGTTGCAACG 768
QY 181 GACAAAGGTTTAAATTAACAGAGCTCATCCAGCATTTAGGAAACATTAAGTTGACATTGAT 240
Db 769 GACAAAGGTTTAAATTAACAGAGCTCATCCAGCATTTAGGAAACATTAAGTTGACATTGAT 828
QY 241 TGCAGTCAGAAAGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
Db 829 TGCAGTCAGAAAGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 888
QY 301 GACTGTTTTCTCACTGTAATCTTGTGGTAACTGGAAGAGCTTCACTGTTTCACTCG 360
Db 889 GACTGTTTTCTCACTGTAATCTTGTGGTAACTGGAAGAGCTTCACTGTTTCACTCG 948
QY 361 CTGACAGGTGGGTTTGTGAGATCTCTGTTGTTTCCCTGGGCTCTGTGAGAGAAATCC 420
Db 949 CTGACAGGTGGGTTTGTGAGATCTCTGTTGTTTCCCTGGGCTCTGTGAGAGAAATCC 1008
QY 421 ACTCTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACCTGTGGCAACATTTGGGCAACC 480
Db 1009 ACTCTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACCTGTGGCAACATTTGGGCAACC 1068
QY 481 CGAATTTCTCCCAATCTTTATCTTGTGCTGCCAGAGATGTTCTTCAACAGAGCTGATA 540
Db 1069 CGAATTTCTCCCAATCTTTATCTTGTGCTGCCAGAGATGTTCTTCAACAGAGCTGATA 1128
QY 541 CAGCAGAAATGGGATTTGGTTATGTGTTAAATCCAGACTATACCTGTCCAAAGCTGACTTT 600
Db 1129 CAGCAGAAATGGGATTTGGTTATGTGTTAAATCCAGACTATACCTGTCCAAAGCTGACTTT 1188
QY 601 ATCCCGAGTCTCAATTTCTGCTGCTGCTGTGATGACAGCTTTTGTGAGAAATTTTG 660
Db 1189 ATCCCGAGTCTCAATTTCTGCTGCTGCTGTGATGACAGCTTTTGTGAGAAATTTTG 1248
QY 661 CCGTGTGTGACCAATATGATGATTTTATTTGAGAAACCAAAAGCTTCAATGATGTGTT 720
Db 1249 CCGTGTGTGACCAATATGATGATTTTATTTGAGAAACCAAAAGCTTCAATGATGTGTT 1308
QY 721 CTAATGCACTGTTTAAAGTGGGATCTCCGCTCCGCAACATGCTATCGCTATCATATG 780
Db 1309 CTAATGCACTGTTTAAAGTGGGATCTCCGCTCCGCAACATGCTATCGCTATCATATG 1368
QY 781 AAGAGATGACATGTTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTACT 840
Db 1369 AAGAGATGACATGTTTATGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGCTTACT 1428
QY 841 ATATCTCCAAATCTTCAATTTTGTGGGCACTCTGACATATGAAAGAAAGATTTAGAAC 900
Db 1429 ATATCTCCAAATCTTCAATTTTGTGGGCACTCTGACATATGAAAGAAAGATTTAGAAC 1488
QY 901 CAGACTGAGACATCAGGAGCCAAAGACAACTCAAGTGTGCACTGAGAGAGCCAAAT 960
Db 1489 CAGACTGAGACATCAGGAGCCAAAGACAACTCAAGTGTGCACTGAGAGAGCCAAAT 1548
QY 961 GAACCTGTCCCTGCTGTCTGAGAGGTGACAGAAAAAGCGCAAGCGCCCTCACTCAACC 1020
Db 1549 GAACCTGTCCCTGCTGTCTGAGAGGTGACAGAAAAAGCGCAAGCGCCCTCACTCAACC 1608
QY 1021 TGTGCGGACCTCTGCTACTCTAGAGGACAGAGCAAAAGGCGCTGCATCCCGCAAGCTG 1080
Db 1609 TGTGCGGACCTCTGCTACTCTAGAGGACAGAGCAAAAGGCGCTGCATCCCGCAAGCTG 1668

QY 1081 CCCAGCGTCCAGCGCTGACGCGCTGCTGTAGAGGACGCGCGTGGTACAGCGCTC 1140
DB 1669 CCCAGCGTCCAGCGCTGACGCGCTGCTGTAGAGGACGCGCGTGGTACAGCGCTC 1728
QY 1141 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCGTTCCTTC 1200
DB 1729 AGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCGTTCCTTC 1788
QY 1201 TCTCTGGATTAACAATCAAGTTTCAATATTCAGCCAGACATGCGACATCTCTTAATGCTTC 1260
DB 1789 TCTCTGGATTAACAATCAAGTTTCAATATTCAGCCAGACATGCGACATCTCTTAATGCTTC 1848
QY 1261 TCCTCATCAGAAAGATGCTTGAATTAACAACCTTCCACTCTGATGGAGCAAC 1320
DB 1849 TCCTCATCAGAAAGATGCTTGAATTAACAACCTTCCACTCTGATGGAGCAAC 1908
QY 1321 AAGCTATGCCAGTTCTCCCTGTTACAGAACTATCGAGAGAGACTCCGAAACAGTCTCT 1380
DB 1909 AAGCTATGCCAGTTCTCCCTGTTACAGAACTATCGAGAGAGACTCCGAAACAGTCTCT 1968
QY 1381 GATTAAGAGAGAACAGCATCCCAAGAGCTGACAGCCGCGCTTACAGACGAC 1440
DB 1969 GATTAAGAGAGAACAGCATCCCAAGAGCTGACAGCCGCGCTTACAGACGAC 2028
QY 1441 AGCAGGATTTGATTCGGTACAGACGAGGAGGAGCGCGCGAGGTCCTTTTA 1500
DB 2029 AGCAGGATTTGATTCGGTACAGACGAGGAGGAGCGCGCGAGGTCCTTTTA 2088
QY 1501 TCTCCACTGATGAGAGTGGGAGCGTGGAGCAATTAACAACAGCTTCCTTTTCGGG 1560
DB 2089 TCTCCACTGATGAGAGTGGGAGCGTGGAGCAATTAACAACAGCTTCCTTTTCGGG 2148
QY 1561 CTTTCCACGACGACGACGACCTTCAAGAGTCTGCTGCGCTTGAAGGCTGCGAC 1620
DB 2149 CTTTCCACGACGACGACGACCTTCAAGAGTCTGCTGCGCTTGAAGGCTGCGAC 2208
QY 1621 TCGGATATCTTGGCCCCCGGAGACTGACCTTCCCTGACGAGAGCTGATTTTGGC 1680
DB 2209 TCGGATATCTTGGCCCCCGGAGACTGACCTTCCCTGACGAGAGCTGATTTTGGC 2268
QY 1681 ACAGAGTCTTCACTTCTACTGCTCAGCCATCTACGAGAGGAGTGCAGTACTCT 1740
DB 2269 ACAGAGTCTTCACTTCTACTGCTCAGCCATCTACGAGAGGAGTGCAGTACTCT 2328
QY 1741 GCTTACAGCTGACGACGACTGCCCACTTGGCGAGACCAAGCTATTTCTGTGCGACGCG 1800
DB 2329 GCTTACAGCTGACGACGACTGCCCACTTGGCGAGACCAAGCTATTTCTGTGCGACGCG 2388
QY 1801 CAGAAGCCAGTGAAGAGTGAAGTGGCGGAGGAGTGGAGTGAAGAGGCGCTTTGA 1860
DB 2389 CAGAAGCCAGTGAAGAGTGAAGTGGCGGAGGAGTGGAGTGAAGAGGCGCTTTGA 2448
QY 1861 AAGCAGTTTAAAGCAGAACTCCAAATGAAATTTGAGAGAGCATCATGTACAGAAAC 1920
DB 2449 AAGCAGTTTAAAGCAGAACTCCAAATGAAATTTGAGAGAGCATCATGTACAGAAAC 2508
QY 1921 AGGTACGAGGAGAGTGGGGAAGTGGGAGTCAAGTCTGATTTTGGGCGACATGGA 1980
DB 2509 AGGTACGAGGAGAGTGGGGAAGTGGGAGTCAAGTCTGATTTTGGGCGACATGGA 2568
QY 1981 ATCATTTAGAGTCTCC 1995
DB 2569 ATCATTTAGAGTCTCC 2583

RESULT 14

ABV21312

ID ABV21312 standard; cDNA; 5145 BP.

XX AC ABV21312;

XX DT 13-SEP-2002 (first entry)

XX

DE Human prostate expression marker cDNA 21303.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW Pharmacogenomic marker; gene; ss.
XX Homo sapiens.
OS MO200160860-A2.
PN 23-AUG-2001.
PD 20-FEB-2001; 2001WO-US005171.
PF 17-FEB-2000; 2000US-0183119P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX MPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3539; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
XX
XX Query Match 94.9%; Score 1893; DB 5; Length 5145;
XX Best Local Similarity 99.9%; Pred No. 0;
XX Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATTGGAATCAATTTGTAAGAGGTTGGCTGTGCGAA 60
DB 589 ATGGCCCATGAGATGATTGGAATCAATTTGTAAGAGGTTGGCTGTGCGAA 648
QY 61 AGTGAACGGAAAAAGTGTCTTAATTGATAGCCGCGCAATTTGTGAATTAACATATCC 120
DB 649 AGTGAACGGAAAAAGTGTCTTAATTGATAGCCGCGCAATTTGTGAATTAACATATCC 708
QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAGGAAAGTTGCAACAG 180
DB 709 CACATTTTGAAGCCATTATATCACTGCTCCAGCTTATGAGGAAAGTTGCAACAG 768
QY 181 GACAAAGTTTATTAACAGAGCTCATCCAGATTACGCGAAACATTAAGTTGATGAT 240
DB 769 GACAAAGTTTATTAACAGAGCTCATCCAGATTACGCGAAACATTAAGTTGATGAT 828
QY 241 TGCAGTCAGAAAGTTGATTTACGATCAAAAGTCCCAAGAGTTGCTCTCTCTTCA 300
DB 829 TGCAGTCAGAAAGTTGATTTACGATCAAAAGTCCCAAGAGTTGCTCTCTCTTCA 888
QY 301 GACTGTTTCTACGTACTCTTGGGTAAATGAGAAAGAGCTCAACTCTGTTACCTG 360

Db 889 GACTGTTTCTCACTGTAATCTTGTGGTAACTGGAGAAAGAGCTTCAACTCTGTTCACTG 948
 QY 361 CTTGACAGGTGGGTTTGTGAGTCTCTGCTGTTTCCCTGGCCCTCTGTGAAGAAATCC 420
 Db 949 CTTGACAGGTGGGTTTGTGAGTCTCTGCTGTTTCCCTGGCCCTCTGTGAAGAAATCC 1008
 QY 421 ACTGTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTTACCTGTGCAACATTTGGGCCAAC 480
 Db 1009 ACTGTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTTACCTGTGCAACATTTGGGCCAAC 1068
 QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCCAGGAGATGCTCTCAACAAAGAGCTGATA 540
 Db 1069 CGAATTTCTCCCAATCTTTATCTTGGCTGCCAGGAGATGCTCTCAACAAAGAGCTGATG 1128
 QY 541 CAGCAGATGGGATTTGGTTATGTGTAAATGCCAGCTATACCTGTCCAAAGCCGTGACTT 600
 Db 1129 CAGCAGATGGGATTTGGTTATGTGTAAATGCCAGCTATACCTGTCCAAAGCCGTGACTT 1188
 QY 601 ATCCCGAGTCTCAATTTCTGCGCTGTGCTGTGAATGACAGCTTTGTGAAAAATTTTG 660
 Db 1189 ATCCCGAGTCTCAATTTCTGCGCTGTGCTGTGAATGACAGCTTTGTGAAAAATTTTG 1248
 QY 661 CCGTGTGTGACAAATAGTAAATTTGATTTGAGAAAGCAAAAGCTTCAATGATGTGT 720
 Db 1249 CCGTGTGTGACAAATAGTAAATTTGATTTGAGAAAGCAAAAGCTTCAATGATGTGT 1308
 QY 721 CTAGTGACGTGTTAGCTGGGATCTCCGCTCCGCCACATCGCTATCGCTTACATCATG 780
 Db 1309 CTAGTGACGTGTTAGCTGGGATCTCCGCTCCGCCACATCGCTTACATCATG 1368
 QY 781 AAGAGATGACATGCTTTTAAATGAATGAAGCTTACAGATTTGTGAAAAAAGAAAGAAAGCTTACT 840
 Db 1369 AAGAGATGACATGCTTTTAAATGAATGAAGCTTACAGATTTGTGAAAAAAGAAAGAAAGCTTACT 1428
 QY 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGTGACATGATGAGAAAGAAAGTTAAGAC 900
 Db 1429 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTGTGACATGATGAGAAAGAAAGTTAAGAC 1488
 QY 901 CAGACTGGAGCATCAGGGCCAAAGCAAACTCAAGCTGTGACCTGTGAGAAAGCCAAAT 960
 Db 1489 CAGACTGGAGCATCAGGGCCAAAGCAAACTCAAGCTGTGACCTGTGAGAAAGCCAAAT 1548
 QY 961 GAACCTGTCCCTGCTGCTCTCAAGGGGTGACAGAAAGCAAGCCCTCACTCAGTCAACC 1020
 Db 1549 GAACCTGTCCCTGCTGCTCTCAAGGGGTGACAGAAAGCAAGCCCTCACTCAGTCAACC 1608
 QY 1021 TGTGCCGACTGTGCTACTCAGAGGACAGAGCAAAAGGCCGTGTATCCGCCAGCGTG 1080
 Db 1609 TGTGCCGACTGTGCTACTCAGAGGACAGAGCAAAAGGCCGTGTATCCGCCAGCGTG 1668
 QY 1081 CCCAGCGTCCCAAGCGTGTGACCGCTGTGTAGAGACAGCCCGCTGTATCAGGCGCTC 1140
 Db 1669 CCCAGCGTCCCAAGCGTGTGACCGCTGTGTAGAGACAGCCCGCTGTATCAGGCGCTC 1728
 QY 1141 AAGGGGCTGACCTGTGCGGCAACAGAGCTGTGAAGACAGCAATTAAGCTCAAGGTTCTTC 1200
 Db 1729 AAGGGGCTGACCTGTGCGGCAACAGAGCTGTGAAGACAGCAATTAAGCTCAAGGTTCTTC 1788
 QY 1201 TCTCTGATATCAAAATCAAGTTTCAATTTAGCCAGAGATGGAGCATCTTTCATGAGCTTC 1260
 Db 1789 TCTCTGATATCAAAATCAAGTTTCAATTTAGCCAGAGATGGAGCATCTTTCATGAGCTTC 1848
 QY 1261 TCTCTATCAAGAAATGCTTTGGAATTAATAAACTTCCACTACTCTGTGATGGAGCAAC 1320
 Db 1849 TCTCTATCAAGAAATGCTTTGGAATTAATAAACTTCCACTACTCTGTGATGGAGCAAC 1908
 QY 1321 AAGCTATGACAGTCTCCCTGTGTGAGAACTATGGAGACAGATCTCCGAAACCAAGTCTCT 1380
 Db 1909 AAGCTATGACAGTCTCCCTGTGTGAGAACTATGGAGACAGATCTCCGAAACCAAGTCTCT 1968
 QY 1381 GATAAGAGAAAGCCAGATCCCAAGAAAGCTGACAGCCGACGAGCTTTCAGACGACAG 1440
 Db 1969 GATAAGAGAAAGCCAGATCCCAAGAAAGCTGACAGCCGACGAGCTTTCAGACGACAG 2028

QY 1441 AGCAAGCATTCATTCCTGCTCAGAACACAGACAGTGGACCGCCAGAGTCCCTTTTA 1500
 Db 2029 AGCAAGCATTCATTCCTGCTCAGAACACAGACAGTGGACCGCCAGAGTCCCTTTTA 2088
 QY 1501 TCTCCACATGCAATGAGTGGAGCGGTGAGGCAATTAACAACAGACTTCTTTTGGC 1560
 Db 2089 TCTCCACATGCAATGAGTGGAGCGGTGAGGCAATTAACAACAGACTTCTTTTGGC 2148
 QY 1561 CTTTCACACAGCCAGACAGACCTCAAGAAATGCTGTGCTGTGGCTTTAAGGGCTGGAC 1620
 Db 2149 CTTTCACACAGCCAGACAGACCTCAAGAAATGCTGTGCTGTGGCTTTAAGGGCTGGAC 2208
 QY 1621 TGGATATCTTGGCCCCCAGACCTTACCCCTTCTGTACACAGAGCTGTATTTGCC 1680
 Db 2209 TGGATATCTTGGCCCCCAGACCTTACCCCTTCTGTACACAGAGCTGTATTTGCC 2268
 QY 1681 ACAGAGTCTCACTTCACTGTGCTCAGCAATCTACGAGAGGAGTGGCAATTAATCT 1740
 Db 2269 ACAGAGTCTCACTTCACTGTGCTCAGCAATCTACGAGAGGAGTGGCAATTAATCT 2328
 QY 1741 GCTTACAGCTGACAGCAGCTGCCCACTTTCGAGAACCAAGTCTATTTGTGCGAGCGG 1800
 Db 2329 GCTTACAGCTGACAGCAGCTGCCCACTTTCGAGAACCAAGTCTATTTGTGCGAGCGG 2388
 QY 1801 CAGAAAGCAAGTGAACAGAGCTGACTTGTGGGCGGAGCTGTGAGAGAGCCCTTTGAA 1860
 Db 2389 CAGAAAGCAAGTGAACAGAGCTGACTTGTGGGCGGAGCTGTGAGAGAGCCCTTTGAA 2448
 QY 1861 AAGCATTTAAACGACAGAAAGCTGCCAAATGGAATTTGAGAGAGCATCATGTGAGAAC 1920
 Db 2449 AAGCATTTAAACGACAGAAAGCTGCCAAATGGAATTTGAGAGAGCATCATGTGAGAAC 2508
 QY 1921 AGGTACCGGGAAGAGCTGGGGAAGAGTGGGAGTCAAGTCACTTTTGGGACAGATGAA 1980
 Db 2509 AGGTACCGGGAAGAGCTGGGGAAGAGTGGGAGTCAAGTCACTTTTGGGACAGATGAA 2568
 QY 1981 ATCATTGAGGTCTCC 1995
 Db 2569 ATCATTGAGGTCTCC 2583

RESULT 15
 ABV21316
 ID ABV21316 standard; cDNA; 5145 BP.
 XX
 AC ABV21316;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 21307.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 OS
 PN WO200160860-A2.
 PN
 PD 23-AUG-2001.
 PD
 PF 20-FEB-2001; 2001WO-US005171.
 PF
 PR 17-FEB-2000; 2000US-0183139P.
 PR 16-MAR-2000; 2000US-0189682P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 PR
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;

Db 2329 GCTTACAGCTGACGACCAAGCTCCCACTTTCGGAGACCAAGTCTATCTGTGCGAGCGG 2388
Qy 1801 CAGAGCCAACTGACAGAGCTGACTCGGCGGAGCTGGCATGAGAGACCCCTTTGAA 1860
Db 2389 CAGAGCCAACTGACAGAGCTGACTCGGCGGAGCTGGCATGAGAGACCCCTTTGAA 2448
Qy 1861 AAGCAGTTTAAACGACAGAGCTGCGCAATGGAATTTGAGAGAGCATCATGTGAGAGAC 1920
Db 2449 AAGCAGTTTAAACGACAGAGCTGCGCAATGGAATTTGAGAGAGCATCATGTGAGAGAC 2508
Qy 1921 AGGTCACGGAGAGAGCTGGGAAAGTGGGAGTCACTCTGCTTTTGGGACAGATGAA 1980
Db 2509 AGGTCACGGAGAGAGCTGGGAAAGTGGGAGTCACTCTGCTTTTGGGACAGATGAA 2568
Qy 1981 ATCATTTGAGGTCTTCC 1995
Db 2569 ATCATTTGAGGTCTTCC 2583

RESULT 16
ABV26826
ID ABV26826 standard; cDNA; 5145 BP.
XX
AC ABV26826;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 26817.
XX
KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183119P.
XX
PR 16-MAR-2000; 2000US-0189862P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer. Useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 5420-5421; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising
a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
specification or its complement. (I) is useful for: (a) assessing whether
a patient is afflicted with prostate cancer; (b) monitoring the
progression of prostate cancer in a patient; (c) assessing the efficacy
of a test compound to inhibit prostate cancer in a patient; (d) assessing
the efficacy of a therapy for inhibiting prostate cancer in a patient;
(e) selecting a composition for inhibiting prostate cancer in a patient;
(f) assessing the prostate cell carcinogenic potential of a compound; (g)
determining whether prostate cancer has metastasized in a patient; (h)
assessing the aggressiveness or indolence of prostate cancer in a patient
; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 94.9%; Score 1893; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCCATGAGATGATTTGGAATCTCAATTTGTACTGAGAGTTGGTCTCTGTGAA 60
Db 589 ATGGCCCATGAGATGATTTGGAATCTCAATTTGTACTGAGAGTTGGTCTCTGTGAA 648
Qy 61 AGTGAACGGAAAAAGTCTGCTAATTTAGATCCGGCAATTTGTGAAATACAAATCATCC 120
Db 649 AGTGAACGGAAAAAGTCTGCTAATTTAGATCCGGCAATTTGTGAAATACAAATCATCC 708
Qy 121 CACATTTTGAAGACCATTAATATCACTCTCCAAAGCTTTATGAAACGAAGTTGCAACAG 180
Db 709 CACATTTTGAAGACCATTAATATCACTCTCCAAAGCTTTATGAAACGAAGTTGCAACAG 768
Qy 181 GACAAAGTTTAAATTAACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 240
Db 769 GACAAAGTTTAAATTAACAGAGCTCATCCAGCATTCAGGAAACATAAGTTGACATTGAT 828
Qy 241 TGCAGTCAAGAGTTTGAATTTAGATCAAAAGCTCCAAAGATTTGCTCTCTTCA 300
Db 829 TGCAGTCAAGAGTTTGAATTTAGATCAAAAGCTCCAAAGATTTGCTCTCTTCA 888
Qy 301 GACTGTTTTCACCTGTACTCTTGGGTAACTGGAGAAAGACTTCAACTCTTCACTTG 360
Db 889 GACTGTTTTCACCTGTACTCTTGGGTAACTGGAGAAAGACTTCAACTCTTCACTTG 948
Qy 361 CTTCAGAGTGGGTTTGGCTGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAAGAAATTC 420
Db 949 CTTCAGAGTGGGTTTGGCTGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAAGAAATTC 1008
Qy 421 ACTCTAGTCCCTACCTGATTTCTACGCTTGTCTTACCTGTGCAACATTGGGCAACC 480
Db 1009 ACTCTAGTCCCTACCTGATTTCTACGCTTGTCTTACCTGTGCAACATTGGGCAACC 1068
Qy 481 CGAATTCCTCCCAATCTTTAATCTTGGCTGCGCAGGAGAAATGCTCTCAACAGAGCTATA 540
Db 1069 CGAATTCCTCCCAATCTTTAATCTTGGCTGCGCAGGAGAAATGCTCTCAACAGAGCTATG 1128
Qy 541 CAGCAGATGGAGATTTGTTATGTTTAAATGCCAGTATACCTGTCAAGAGCTGACTTT 600
Db 1129 CAGCAGATGGAGATTTGTTATGTTTAAATGCCAGTATACCTGTCAAGAGCTGACTTT 1188
Qy 601 ATCCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG 660
Db 1189 ATCCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAAAATTTTG 1248
Qy 661 CCGTGTGGAACAATCAGTATGATTTCAATGAGAAAGCAAAAGCCCTCAATGATGTGT 720
Db 1249 CCGTGTGGAACAATCAGTATGATTTCAATGAGAAAGCAAAAGCCCTCAATGATGTGT 1308
Qy 721 CTAGTGACATGTTTAAAGTGGATCTCCGCTCCGCCACATCGCTATGCTCATCATG 780
Db 1309 CTAGTGACATGTTTAAAGTGGATCTCCGCTCCGCCACATCGCTATGCTCATCATG 1368
Qy 781 AAGAGATGACATGCTTTTAATGAAGCTTACAGATTTGTGAAAAAAGAAAGCACTTACT 840
Db 1369 AAGAGATGACATGCTTTTAATGAAGCTTACAGATTTGTGAAAAAAGAAAGCACTTACT 1428
Qy 841 ATATCTCCAAACTTCAATTTTCTGGGCAATCTGAGCATATGAGAAAGAAATTAAGAC 900
Db 1429 ATATCTCCAAACTTCAATTTTCTGGGCAATCTGAGCATATGAGAAAGAAATTAAGAC 1488
Qy 901 CAGACTGAGACATCAGAGGCCAAAGAGCAAACTCAAGCTCTGACCTGAGAGAACCAAT 960
Db 1489 CAGACTGAGACATCAGAGGCCAAAGAGCAAACTCAAGCTCTGACCTGAGAGAACCAAT 1548
Qy 961 GAACCTGTCTCTGCTCTCAAGAGGTGACAGAAAGGAGACGCGCTCAGTCCACCC 1020
Db 1549 GAACCTGTCTCTGCTCTCAAGAGGTGACAGAAAGGAGACGCGCTCAGTCCACCC 1608

QY 1021 TGTGGCACTGTCTACTCTCAGAGGAGCAGACGAAAGCCCGTGCATCCCGCAGCGTG 1080
DB 1609 TGTGGCACTGTCTACTCTCAGAGGAGCAGACGAAAGCCCGTGCATCCCGCAGCGTG 1668
QY 1081 CCCAGCGTCCGACGCGTGCAGCGGTGTGTAGAGAGACCCCGCTGTACAGCGCGTC 1140
DB 1669 CCCAGCGTCCGACGCGTGCAGCGGTGTGTAGAGAGACCCCGCTGTACAGCGCGTC 1728
QY 1141 AGTGGGTGCACTGTCCGAGCAGCGCTGGAAGACAGCAATTAAGCTCAAGCGCTTCCTTC 1200
DB 1729 AGTGGGTGCACTGTCCGAGCAGCGCTGGAAGACAGCAATTAAGCTCAAGCGCTTCCTTC 1788
QY 1201 TCTCTGATATCAAAATCAGTTTATTAATTCAGCCAGCATGCGAGCATCTTACATGCGCTTC 1260
DB 1789 TCTCTGATATCAAAATCAGTTTATTAATTCAGCCAGCATGCGAGCATCTTACATGCGCTTC 1848
QY 1261 TCTCTCAGAGAGATGTTTGAATTAATACTCAAACTTTCATCTGTGATGGAGCCAC 1320
DB 1849 TCTCTCAGAGAGATGTTTGAATTAATACTCAAACTTTCATCTGTGATGGAGCCAC 1908
QY 1321 AAGCTATGCCAGTTTCCCTGTTCAGAGAACTATCGAGAGACACTCCGAAACCAAGTCTCT 1380
DB 1909 AAGCTATGCCAGTTTCCCTGTTCAGAGAACTATCGAGAGACACTCCGAAACCAAGTCTCT 1968
QY 1381 GATAAGAGAGAGCCAGCATCCCAAGAGCTGAGACCGCCAGGCTTCAGACAGCAG 1440
DB 1969 GATAAGAGAGAGCCAGCATCCCAAGAGCTGAGACCGCCAGGCTTCAGACAGCAG 2028
QY 1441 AGCAGCGATTTGATTCGATCAGAACCCAGCAGCAGCGCCAGAGGTCCCTTTTA 1500
DB 2029 AGCAGCGATTTGATTCGATCAGAACCCAGCAGCAGCGCCAGAGGTCCCTTTTA 2088
QY 1501 TCTCCACTGATGGAATGGAGCGCTGAGAGACATTAACAACCACTTCTTTTCGCG 1560
DB 2089 TCTCCACTGATGGAATGGAGCGCTGAGAGACATTAACAACCACTTCTTTTCGCG 2148
QY 1561 CTTTCACACGACGACGACGACCTCAAGAGTGTGCTGAGCGCTTAAAGGCTGAGC 1620
DB 2149 CTTTCACACGACGACGACGACCTCAAGAGTGTGCTGAGCGCTTAAAGGCTGAGC 2208
QY 1621 TCGGATATCTTGGCCCCCAGACCTCTAACCCTTCCCTGACAGCAGCTGTATTTTGGC 1680
DB 2209 TCGGATATCTTGGCCCCCAGACCTCTAACCCTTCCCTGACAGCAGCTGTATTTTGGC 2268
QY 1681 ACGAGTCTCTCAACTTCTACTCTGCTCAGCCATCTAGCGAGGACAGTGTCACTCT 1740
DB 2269 ACGAGTCTCTCAACTTCTACTCTGCTCAGCCATCTAGCGAGGACAGTGTCACTCT 2328
QY 1741 GCCTACAGCTGACGACGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGGCGCAGCGG 1800
DB 2329 GCCTACAGCTGACGACGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGGCGCAGCGG 2388
QY 1801 CAGAGCCCAAGTACAGAGCTGACTCGCGCGAGCTGATGAAGAGAGCCCTTTGAA 1860
DB 2389 CAGAGCCCAAGTACAGAGCTGACTCGCGCGAGCTGATGAAGAGAGCCCTTTGAA 2448
QY 1861 AAGCAGTTTAAAGCAGAGCTGCCCAATGAAATTTGAGAGAGCATCATGTGAGAGAAC 1920
DB 2449 AAGCAGTTTAAAGCAGAGCTGCCCAATGAAATTTGAGAGAGCATCATGTGAGAGAAC 2508
QY 1921 AGGTACCGGGAAGAGCTGGGGAAGTGGGAGTCAAGTCTTTTCGGGCAACATGGA 1980
DB 2509 AGGTACCGGGAAGAGCTGGGGAAGTGGGAGTCAAGTCTTTTCGGGCAACATGGA 2568
QY 1981 ATCATTGAGGTCTCC 1995
DB 2569 ATCATTGAGGTCTCC 2583

RESULT 17
ABV27131
ID ABV27131 standard; cDNA; 5145 BP.
XX

AC ABV27131;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 27122.
XX
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN MO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183119P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer. Stage of prostate cancer.
XX
XX Claim 1; Page 5499-5500; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (i) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement; (i) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
Query Match 94.9%; Score 1893; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGAGATGATTGGAACCTCAATTTTACTGAGAGTGTGCTGCTGGA 60
DB 589 ATGGCCCATGAGATGATTGGAACCTCAATTTTACTGAGAGTGTGCTGCTGGA 648
QY 61 AGTGAAGGGAAGAAAGTGTGCTAAATGATGAGCGGCGCAATTTGGAATACATATCC 120
DB 649 AGTGAAGGGAAGAAAGTGTGCTAAATGATGAGCGGCGCAATTTGGAATACATATCC 708
QY 121 CACATTTTGAAGCCATTATATCACTGCTCCAACTTATGAGGGAAGTTGCAACAG 180
DB 709 CACATTTTGAAGCCATTATATCACTGCTCCAACTTATGAGGGAAGTTGCAACAG 768
QY 181 GACAAAGTTTAAATTAACAAGCTCATCCGCAATTCAGGCAACATTAAGTTGACATTGAT 240
DB 769 GACAAAGTTTAAATTAACAAGCTCATCCGCAATTCAGGCAACATTAAGTTGACATTGAT 828
QY 241 TCGAGTCAGAGGTTGATTTACGATCAAGAGTCCCAAGATGTTGCTCTGCTTCA 300

Db 829 TCGACTCAGAAAGTTGATGTATACGATCAAAAGCTCCAAAGATGTTCCTCTCTCTCA 888
QY 301 GACTGTTTCTCAGTGACTTCTGAGTAACTGAGAGAGACTTCAATCTGTTCACTG 360
Db 889 GACTGTTTCTCAGTGACTTCTGAGTAACTGAGAGAGACTTCAATCTGTTCACTG 948
QY 361 CTTCGAGGTGGGTTTGTGAGTCTCTGTTGTTTCCCTGGCCCTCTGGAAGAAAATCC 420
Db 949 CTTCGAGGTGGGTTTGTGAGTCTCTGTTGTTTCCCTGGCCCTCTGGAAGAAAATCC 1008
QY 421 ACTGAGTCCCTACCTGATTTCTCAGGCTTCACTCTGTTGCCAATTTGGGCCAAC 480
Db 1009 ACTGAGTCCCTACCTGATTTCTCAGGCTTCACTCTGTTGCCAATTTGGGCCAAC 1068
QY 481 CGAATTTCTCCCAATCTTATCTTGCTGCGCAGGAGATGTCCCAACAGAGGCTGATA 540
Db 1069 CGAATTTCTCCCAATCTTATCTTGCTGCGCAGGAGATGTCCCAACAGAGGCTGATA 1128
QY 541 CAGCAGAAATGGGATTTGGTATGTGTAAATGCCAGCTAATCCTGTCCAAAGCCTGACTT 600
Db 1129 CAGCAGAAATGGGATTTGGTATGTGTAAATGCCAGCAATACCTGTCCAAAGCCTGACTT 1188
QY 601 ATCCCGAGTCTCATTTCCGCGGTGCTGATGAATGACAGCTTTTGTGAAAATTTTGG 660
Db 1189 ATCCCGAGTCTCATTTCCGCGGTGCTGATGAATGACAGCTTTTGTGAAAATTTTGG 1248
QY 661 CCGTGTGTGACAAATCAGTAGATTTTCAATGAGAAACAAAGCCTTCAATGATGTGT 720
Db 1249 CCGTGTGTGACAAATCAGTAGATTTTCAATGAGAAACAAAGCCTTCAATGATGTGT 1308
QY 721 CTAGTGAATGTTAGTGGGATCTCCCGCTCCGCCACATGCTATGCTTACATCATG 780
Db 1309 CTAGTGAATGTTAGTGGGATCTCCCGCTCCGCCACATGCTTACATCATG 1368
QY 781 AAGAGATGACATGTCTTATAGTAAAGCTTACAGATTTGTGAAAAGAAAAGCACTACT 840
Db 1369 AAGAGATGACATGTCTTATAGTAAAGCTTACAGATTTGTGAAAAGAAAAGCACTACT 1428
QY 841 ATATCTCCAAATCTCAATTTTGTGGGCCAACTCTGAGATGAGAAAGATTTAAGAAC 900
Db 1429 ATATCTCCAAATCTCAATTTTGTGGGCCAACTCTGAGATGAGAAAGATTTAAGAAC 1488
QY 901 CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAACTGCTGACCTGAGAGAACCAAT 960
Db 1489 CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAACTGCTGACCTGAGAGAACCAAT 1548
QY 961 GAACTGTCTCTGCTCTCAGAGGTTGACAGAAAAGGAGACGCCCTCAGTCAACCC 1020
Db 1549 GAACTGTCTCTGCTCTCAGAGGTTGACAGAAAAGGAGACGCCCTCAGTCAACCC 1608
QY 1021 TGTGCGGACTTGTCTTCAAGAGGACAGCAAAAGGCCGTGATCCGCCAGCGTG 1080
Db 1609 TGTGCGGACTTGTCTTCAAGAGGACAGCAAAAGGCCGTGATCCGCCAGCGTG 1668
QY 1081 CCCAGGTCGCCAGCGTGCAGCGGTGTAGAGGACAGGCCCTGTACAGGCGTGC 1140
Db 1669 CCCAGGTCGCCAGCGTGCAGCGGTGTAGAGGACAGGCCCTGTACAGGCGTGC 1728
QY 1141 AGTGGGCTGCACCTGTCCGACAGCGCTGAAAGCAGCAATTAAGCTTCTTTC 1200
Db 1729 AGTGGGCTGCACCTGTCCGACAGCGCTGAAAGCAGCAATTAAGCTTCTTTC 1788
QY 1201 TCTCGGATATCAAAATCAGTTTCAATTCAGCCAGATGGCAGCATCTTCAATGGCTTC 1260
Db 1789 TCTCGGATATCAAAATCAGTTTCAATTCAGCCAGATGGCAGCATCTTCAATGGCTTC 1848
QY 1261 TCCCTATCAGAAAGATGCTTTGGAATCTAACAACCTTCACTACTGAGTGGGACCAAC 1320
Db 1849 TCCCTATCAGAAAGATGCTTTGGAATCTAACAACCTTCACTACTGAGTGGGACCAAC 1908
QY 1321 AAGCATATGCCAGTTCTCCCTGTTTCAAGAACTATCGAGCAGACTCCGAAACCAAGTCTT 1380
Db 1909 AAGCATATGCCAGTTCTCCCTGTTTCAAGAACTATCGAGCAGACTCCGAAACCAAGTCTT 1968

QY 1381 GATTAAGAGAAACCCAGCATTCCTCAAGAAAGCTGACAGACCGCCAGCTTACAGACGCCAG 1440
Db 1969 GATTAAGAGAAAGCCAGCATTCCTCAAGAAAGCTGACAGACCGCCAGCTTACAGACGCCAG 2028
QY 1441 AGCAGAGATTCGATTCGGTCAAGAACCCAGACATGTGGACCGCCAGAGGTCCCTTTTA 1500
Db 2029 AGCAGAGATTCGATTCGGTCAAGAACCCAGACATGTGGACCGCCAGAGGTCCCTTTTA 2088
QY 1501 TCTCCACTGCAATGAAATGGGAGCGTGGAGGACAAATTAACAACAGCAGCTTCTTTTCGAC 1560
Db 2089 TCTCCACTGCAATGAAATGGGAGCGTGGAGGACAAATTAACAACAGCAGCTTCTTTTCGAC 2148
QY 1561 CTTTCACACAGCCAGCAGACACTTCAAGAAATCTGTGCTGGGCTTAAAGGCTGGAC 1620
Db 2149 CTTTCACACAGCCAGCAGACACTTCAAGAAATCTGTGCTGGGCTTAAAGGCTGGAC 2208
QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGCAGCTGTATTTTGGC 1680
Db 2209 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGCAGCTGTATTTTGGC 2268
QY 1681 ACAGAGTCTCAGACTTCTACTCTGCTCAGCCATCTTACGAGGACAGTCCAGTTACTCT 1740
Db 2269 ACAGAGTCTCAGACTTCTACTCTGCTCAGCCATCTTACGAGGACAGTCCAGTTACTCT 2328
QY 1741 GCTTACAGCTCAGCAGCAGCTGCTTCCAGGAGACCAAGTCTTATCTGTGCGACGGCG 1800
Db 2329 GCTTACAGCTCAGCAGCAGCTGCTTCCAGGAGACCAAGTCTTATCTGTGCGACGGCG 2388
QY 1801 CAGAAAGCAAGTGAACAGAGTGAATCTCGCGGCGAGCTGGCATGAAGAGAGGCCCTTTGAA 1860
Db 2389 CAGAAAGCAAGTGAACAGAGTGAATCTCGCGGCGAGCTGGCATGAAGAGAGGCCCTTTGAA 2448
QY 1861 AAGCATTTAAACCCAGAAAGCTGCCAAATGGAATTTGAGAGAGACATATGTACAGAAC 1920
Db 2449 AAGCATTTAAACCCAGAAAGCTGCCAAATGGAATTTGAGAGAGACATATGTACAGAAC 2508
QY 1921 AGTCAAGGAGAGAGCTGGGGAAGTGGGAGTCAAGTCTGCTTTGGGCAAGATGAA 1980
Db 2509 AGTCAAGGAGAGAGCTGGGGAAGTGGGAGTCAAGTCTGCTTTGGGCAAGATGAA 2568
QY 1981 ATCATTTGAGTCTCC 1995
Db 2569 ATCATTTGAGTCTCC 2583

RESULT 18
ABV26923
ID ABV26923 standard; cDNA; 5145 BP.
XX
AC ABV26923;
XX
DT 16-SEP-2002 (first entry)
XX
Human prostate expression marker cDNA 26914.
DB
Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200160860-A2.
PD
XX 23-AUG-2001.
PF 20-FEB-2001; 2001MO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
XX
PR 16-MAR-2000; 2000US-0189862P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA Schlegel R, Endege WO, Monahan JR;
PI MPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 5449-5450; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABY0010-ABY62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
SQ
Query Match 94.9%; Score 1893; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCCATGATGATGGAATCAATGTTACTGAGAGGTTGGTCTGCTGGAA 60
DB 589 ATGGCCCATGATGATGGAATCAATGTTACTGAGAGGTTGGTCTGCTGGAA 648
QY 61 AGTGAACCGAAGAGGCTGCTAATGATGAGCGGCAATTTGGATCAATATAC 120
DB 649 AGTGAACCGAAGAGGCTGCTAATGATGAGCGGCAATTTGGATCAATATAC 708
QY 121 CACATTTTGGAGACATTAATATCACTGCTCAAGCTTATGAAGGAGTTGCAAC 180
DB 709 CACATTTTGGAGACATTAATATCACTGCTCAAGCTTATGAAGGAGTTGCAAC 768
QY 181 GACAAAGTGTAAATACAGAGCTCATCAGCAATTCAGCGAACAATAGTTGACAT 240
DB 769 GACAAAGTGTAAATACAGAGCTCATCAGCAATTCAGCGAACAATAGTTGACAT 828
QY 241 TGGAGTCAGAAAGTTGATTAAGATCAAGATCCCAAGATGTTGCTCTCTTCA 300
DB 829 TGGAGTCAGAAAGTTGATTAAGATCAAGATCCCAAGATGTTGCTCTCTTCA 888
QY 301 GACATTTTCTCACTGACTTCTGGGTTAACTGAGAGAGCTTCAACTGTTTCACT 360
DB 889 GACATTTTCTCACTGACTTCTGGGTTAACTGAGAGAGCTTCAACTGTTTCACT 948
QY 361 CTTCAGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTGTGTGAAGAAATCC 420
DB 949 CTTCAGAGTGGGTTTGTGAGTTCTCTGTTGTTTCCCTGGCTGTGTGAAGAAATCC 1008
QY 421 ACTCTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACCTGTGACCAATTTGGCC 480
DB 1009 ACTCTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACCTGTGACCAATTTGGCC 1068
QY 481 CGAATTTTCCCAATCTTTATCTTGGCTGCAGAGAGATGCTCTCAACAAGAGCTGATA 540
DB 1069 CGAATTTTCCCAATCTTTATCTTGGCTGCAGAGAGATGCTCTCAACAAGAGCTGATA 1128
QY 541 CAGCAGATGGATGGTATGTTATGTTAAATGCCAGCTATACTGTCCAAAGCTGACTTT 600
DB 1129 CAGCAGATGGATGGTATGTTATGTTAAATGCCAGCTATACTGTCCAAAGCTGACTTT 1188
QY 601 ATCCCGAGTCTCATTTCCGCGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 660

DB 1189 ATCCCGAGTCTCATTTCTGCTGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 1248
QY 661 CCGTGGTTGGACAAATCAGTATGATTTTATTGGAAGCAAAAGCTTCCATGATGTTGTT 720
DB 1249 CCGTGGTTGGACAAATCAGTATGATTTTATTGGAAGCAAAAGCTTCCATGATGTTGTT 1308
QY 721 CTAGTGCATGTTTGTGCTGGGATCTCCGCTCCGCGACCATGCTATGCTTACATCATG 780
DB 1309 CTAGTGCATGTTTGTGCTGGGATCTCCGCTCCGCGACCATGCTATGCTTACATCATG 1368
QY 781 AAGAGATGAGATGCTTATGATGAGCTTACAGATTTGTGAAGAAAAAGACCTACT 840
DB 1369 AAGAGATGAGATGCTTATGATGAGCTTACAGATTTGTGAAGAAAAAGACCTACT 1428
QY 841 ATATCTCCAAATTCATATTTTCTGGGCCAATCTCTGAGTATGAGAAATTAAGAAC 900
DB 1429 ATATCTCCAAATTCATATTTTCTGGGCCAATCTCTGAGTATGAGAAATTAAGAAC 1488
QY 901 CAGACTGAGCATCAGGCGCAAGCAAACTCAAGCTGCTGACCTGAGAAAGCCAAAT 960
DB 1489 CAGACTGAGCATCAGGCGCAAGCAAACTCAAGCTGCTGACCTGAGAAAGCCAAAT 1548
QY 961 GAACCTGTCCCTGCTGCTCTCAGAGGTTGAGACAGAAAGCAGAGCCCTCAGTCCACC 1020
DB 1549 GAACCTGTCCCTGCTGCTCTCAGAGGTTGAGACAGAAAGCAGAGCCCTCAGTCCACC 1608
QY 1021 TGTGCCACTGCTTACTTCAAGAGCAGAGCAAAAGCCGCTGATCCCGCAGCGTG 1080
DB 1609 TGTGCCACTGCTTACTTCAAGAGCAGAGCAAAAGCCGCTGATCCCGCAGCGTG 1668
QY 1081 CCCAGCTGCCAGGCTGAGAGCCGCTGTTAGAGAGAGCCGCTGTTACAGAGCGCTC 1140
DB 1669 CCCAGCTGCCAGGCTGAGAGCCGCTGTTAGAGAGAGCCGCTGTTACAGAGCGCTC 1728
QY 1141 AGTGGCTGCACTGCTGCTGCTGAGAGAGCTGAGAGAGCAATTAAGCTCAAGGCTTCTTC 1200
DB 1729 AGTGGCTGCACTGCTGCTGCTGAGAGAGCTGAGAGAGCAATTAAGCTCAAGGCTTCTTC 1788
QY 1201 TCTCTGATATCAATCAATCAATTTTATATTCAGCCAGATGAGCATCTTACATGCTTC 1260
DB 1789 TCTCTGATATCAATCAATCAATTTTATATTCAGCCAGATGAGCATCTTACATGCTTC 1848
QY 1261 TCTCTATCAGAAAGTCTTTGGAATCTAACAACCTTCCATCTCTGATGAGGAGCAAC 1320
DB 1849 TCTCTATCAGAAAGTCTTTGGAATCTAACAACCTTCCATCTCTGATGAGGAGCAAC 1908
QY 1321 AAGCTATGCAATTTCTCCCTGCTTCAAGAACTATGAGAGCAGATCCGAAACAGTCTT 1380
DB 1909 AAGCTATGCAATTTCTCCCTGCTTCAAGAACTATGAGAGCAGATCCGAAACAGTCTT 1968
QY 1381 GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGAGACCGGCAAGCTTCAAGACCCAG 1440
DB 1969 GATTAAGAGAAAGCCAGCATCCCAAGAAAGCTGAGACCGGCAAGCTTCAAGACCCAG 2028
QY 1441 AGCAAGCATTTGATTTGGTCAAGAACACAGACAGTGGCAGAGGCTTCCCTTTTA 1500
DB 2029 AGCAAGCATTTGATTTGGTCAAGAACACAGACAGTGGCAGAGGCTTCCCTTTTA 2088
QY 1501 TCTCCAGTGCATGAGAGTGGAGAGCTGAGAGCAATTAACAACAGCTTCTCTTTCGGC 1560
DB 2089 TCTCCAGTGCATGAGAGTGGAGAGCTGAGAGCAATTAACAACAGCTTCTCTTTCGGC 2148
QY 1561 CTTTCCACAGCCAGCAGACCTCAAGAAATCTGCTGAGCTTGAAGGCTGGCAC 1620
DB 2149 CTTTCCACAGCCAGCAGACCTCAAGAAATCTGCTGAGCTTGAAGGCTGGCAC 2208
QY 1621 TGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGACGCTGTATTTTGGC 1680
DB 2209 TGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGACGCTGTATTTTGGC 2268
QY 1681 ACAGATCTCTCACTTACTCTGCTGAGCATATGAGAGGAGTCCAGTTACTCT 1740

CC	assessing the aggressiveness or indolence of prostate cancer in a patient
CC	; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX	
SQ	Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
	Query Match 94.9%; Score 1893; DB 5; Length 5145;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0
QY	1 ATGGCCCATGAGATGATTGGAACTCAATTGTACTAGAGGTTGGNGCTCGCTGGA 60
DB	589 ATGGCCCATGAGATGATTGGAACTCAATTGTACTAGAGGTTGGNGCTCGCTGGA 648
QY	61 AGTGGAAACGAAAAAGTCGCTCTAATTGATGACCGGCATTTGNGAATAACAATCCTC 120
DB	649 AGTGGAAACGAAAAAGTCGCTCTAATTGATGACCGGCATTTGNGAATAACAATCCTC 708
QY	121 CACATTTTGGAAGCATTAATATACAATGCTCCAAGTTATGAAGCAAGGTTGCCAACG 180
DB	709 CACATTTTGGAAGCATTAATATACAATGCTCCAAGTTATGAAGCAAGGTTGCCAACG 768

•

2

2

३५

Cur

02

2

100
 nh

4807

106

54 0v

Dib 112

0v 60

Db 118

66 OY

D**b** 124

QY 72

Db 130

QY 78

Db 136

842Y

D5 142

90: 2y

1489

Db 769 GACAAAGTGTATTAACAGAGCTCATCCAGCATCCAGGAAACATAGAGTTGACATTGAT 828
 QY 241 TCGAGTCAGAGAGTTGATGTTAGCATGAAAGCTCCCAAGATGTTGCTCTCTCTCTTCA 300
 Db 829 TCGAGTCAGAGAGTTGATGTTAGCATGAAAGCTCCCAAGATGTTGCTCTCTCTCTTCA 888
 QY 301 GACTGTTTTCTCAGTACTTCTGGTAAACTGAGAGAGCTTCAACTCTGTTCACTG 360
 Db 889 GACTGTTTTCTCAGTACTTCTGGTAAACTGAGAGAGCTTCAACTCTGTTCACTG 948
 QY 361 CTGGAGGTGGGTGCTGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAAGAAAATCC 420
 Db 949 CTGGAGGTGGGTGCTGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAAGAAAATCC 1008
 QY 421 ACTTAGTCCCTACCTGATTTCTCAGGCTTGCTTAACCTGTTGCAACATTGGGCCAAC 480
 Db 1009 ACTTAGTCCCTACCTGATTTCTCAGGCTTGCTTAACCTGTTGCAACATTGGGCCAAC 1068
 QY 481 CGAATTCCTCCCAATCTTTATCTTGCTGCTGCCAGGAGATGTCCTCAACAGAGCTGATA 540
 Db 1069 CGAATTCCTCCCAATCTTTATCTTGCTGCTGCCAGGAGATGTCCTCAACAGAGCTGATA 1128
 QY 541 CAGAGAGATGGGATTTGTTATGTTAAATGCCCAGCTATACCTGTCCAAAGCTGACTTT 600
 Db 1129 CAGAGAGATGGGATTTGTTATGTTAAATGCCCAGCAATACCTGTCCAAAGCTGACTTT 1188
 QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGAAATGACAGCTTTGTGAGAAAATTTTG 660
 Db 1189 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGAAATGACAGCTTTGTGAGAAAATTTTG 1248
 QY 661 CCGGTGTTGACAAATCAGTAGATTTATTGAGAGAGAGAAAGCTTCCAAATGAGATGTT 720
 Db 1249 CCGGTGTTGACAAATCAGTAGATTTATTGAGAGAGAGAAAGCTTCCAAATGAGATGTT 1308
 QY 721 CTAGTGAAGTGTAGTGGGATCTCCGCTCCGCAACATGCTATGCCCTTACATCATG 780
 Db 1309 CTAGTGAAGTGTAGTGGGATCTCCGCTCCGCAACATGCTATGCCCTTACATCATG 1368
 QY 781 AAGAGATGACATGTCCTTTAGATGAAGCTTACAGATTTGTGAGAGAGAAAGAGCTTACT 840
 Db 1369 AAGAGATGACATGTCCTTTAGATGAAGCTTACAGATTTGTGAGAGAGAAAGAGCTTACT 1428
 QY 841 ATATCTCCAAACTCAATTTCTGGGCCAACTCTCTGACTATGAGAGAGATTTAAGAAC 900
 Db 1429 ATATCTCCAAACTCAATTTCTGGGCCAACTCTGACTATGAGAGAGATTTAAGAAC 1488
 QY 901 CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAAGCTCTGCACTCGAGAGAGCCAAAT 960
 Db 1489 CAGACTGAGACATCAGGGCCAAAGAGCAAACTCAAGCTCTGCACTCGAGAGAGCCAAAT 1548
 QY 961 GAACTGTCTCTGCTCTCTCAAGGGGTGACAGAAAAGGAGAGCCCTCTCAGTCCACC 1020
 Db 1549 GAACTGTCTCTGCTCTCTCAAGGGGTGACAGAAAAGGAGAGCCCTCTCAGTCCACC 1608
 QY 1021 TGTGCCGACTGTCTCTCAAGGGGTGACAGAAAAGGAGAGCCCTCTCAGTCCACC 1080
 Db 1609 TGTGCCGACTGTCTCTCAAGGGGTGACAGAAAAGGAGAGCCCTCTCAGTCCACC 1668
 QY 1081 CCCAGGCTGCCAGGGGTGACAGCCGTGTTAGAGAGAGCCCTCTGTTACAGGCGCTC 1140
 Db 1669 CCCAGGCTGCCAGGGGTGACAGCCGTGTTAGAGAGAGCCCTCTGTTACAGGCGCTC 1728
 QY 1141 AGTGGGCTGACCTGTCCGAGAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 Db 1729 AGTGGGCTGACCTGTCCGAGAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1788
 QY 1201 TCTCTGATATCAATCAAGTTTCAATTCAGCCAGCATGGAGAGATCCTTACATGAGCTTC 1260
 Db 1789 TCTCTGATATCAATCAAGTTTCAATTCAGCCAGCATGGAGAGATCCTTACATGAGCTTC 1848
 QY 1261 TCTCTATCAAGAGATGCTTTGAGATGACAAACCTTCCACTACTCTGATGGAGAGAGAG 1320
 Db 1849 TCTCTATCAAGAGATGCTTTGAGATGACAAACCTTCCACTACTCTGATGGAGAGAGAG 1908

QY 1321 AAGCTATGCAAGTTCCTCCCTGTGAGGAAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 Db 1909 AAGCTATGCAAGTTCCTCCCTGTGAGGAAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1968
 QY 1381 GATTAAG 1440
 Db 1969 GATTAAG 2028
 QY 1441 AGCAAGCATTCATTCGGTGTGAG 1500
 Db 2029 AGCAAGCATTCATTCGGTGTGAG 2088
 QY 1501 TCTCCACTGATCGAAGTGGAG 1560
 Db 2089 TCTCCACTGATCGAAGTGGAG 2148
 QY 1561 CTTTCACAG 1620
 Db 2149 CTTTCACAG 2208
 QY 1621 TCGATATCTTGGCCGCCAGAGCTTACCCCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
 Db 2209 TCGATATCTTGGCCGCCAGAGCTTACCCCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2268
 QY 1681 ACAAGTCTCTCAACTTCTACTGCTGCTCAGAGATCTTACGAGAGAGAGAGAGAGAGAGAGAGAG 1740
 Db 2269 ACAAGTCTCTCAACTTCTACTGCTGCTCAGAGATCTTACGAGAGAGAGAGAGAGAGAGAGAGAG 2328
 QY 1741 GCTTACAGCTGACAG 1800
 Db 2329 GCTTACAGCTGACAG 2388
 QY 1801 CAGAGCCAAAGTGAAG 1860
 Db 2389 CAGAGCCAAAGTGAAG 2448
 QY 1861 AAGCATTTTAAAGCAG 1920
 Db 2449 AAGCATTTTAAAGCAG 2508
 QY 1921 AGGTCAAGGAG 1980
 Db 2509 AGGTCAAGGAG 2568
 QY 1981 ATCATTGAGGTCTCC 1995
 Db 2569 ATCATTGAGGTCTCC 2583

RESULT 21
 ID ABV22827 standard; cDNA, 5145 BP.
 XX AC ABV22827;
 XX DT 13-SEP-2002 (first entry)
 XX DE Human prostate expression marker cdna 22818.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN MO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US005171.
 XX PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.

Db 2209 TCGATATCTTGGCCCCCAGACCTCTACCCCTTCCTGACCAAGAGCTGATTTGCC 2268
Qy 1681 ACAGAGTCTCACACTTTACTCTGCTCCCTCAGCCATCTACGAGGAGTGGCAATTTACTCT 1740
Db 2269 ACAGAGTCTCACACTTTACTCTGCTCCCTCAGCCATCTACGAGGAGTGGCAATTTACTCT 2328
Qy 1741 GCTTACAGCTCAGCAGCTGCTCCCACTTGGCGAGACCAAGTCTATTCTGTCGCGAGGCGG 1800
Db 2329 GCTTACAGCTCAGCAGCTGCTCCCACTTGGCGAGACCAAGTCTATTCTGTCGCGAGGCGG 2388
Qy 1801 CAGAGCCCAATGACAGAGCTGACTCCGCGCGAGAGTGGCAATGAGAGAGCCCTTTGAA 1860
Db 2389 CAGAGCCCAATGACAGAGCTGACTCCGCGCGAGAGTGGCAATGAGAGAGCCCTTTGAA 2448
Qy 1861 AAGCAGTTTAAACCAAGAGCTGCAATGGAATTTGAGAGAGCATGATGAGAGAAC 1920
Db 2449 AAGCAGTTTAAACCAAGAGCTGCAATGGAATTTGAGAGAGCATGATGAGAGAAC 2508
Qy 1921 AGGTCAAGGAGAGAGCTGGGGGAAAGTGGGAGTCACTGCTTTTGGGCGAGCATGAA 1980
Db 2509 AGGTCAAGGAGAGAGCTGGGGGAAAGTGGGAGTCACTGCTTTTGGGCGAGCATGAA 2568
Qy 1981 ATCATTAAGTCTCC 1995
Db 2569 ATCATTAAGTCTCC 2583

RESULT 22

ABV26934
ID ABV26934 standard; cDNA; 5145 BP.

XX AC ABV26934;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 26925.

XX KM Human; prostate cancer; cytosolic; carcinogen; pharmacodynamic marker;

XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN W0200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of

XX CC prostate cells and correlating with presence of prostate cancer, useful

XX CC for detecting presence of prostate cancer, stage of prostate cancer.

XX PS Claim 1; Page 5453-5454; 11750pp; English.

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
Query Match 94.9%; Score 1893; DB 5; Length 5145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCCCATGAGATATTGGAATCTCAATTTGTAATGAGAGTTGGTCTCTCTGGA 60
Db 589 ATGGCCCATGAGATATTGGAATCTCAATTTGTAATGAGAGTTGGTCTCTCTGGA 648
Qy 61 AGTGAACGGAAAAAGTCTGCTAATTGATGCGCGCATTTTGGAAATACATATCC 120
Db 649 AGTGAACGGAAAAAGTCTGCTAATTGATGCGCGCATTTTGGAAATACATATCC 708
Qy 121 CACATTTTGAAGCCATTATATCACTGCTCCAAAGCTTATGAGCGAAGTTGCAACAG 180
Db 709 CACATTTTGAAGCCATTATATCACTGCTCCAAAGCTTATGAGCGAAGTTGCAACAG 768
Qy 181 GACAAAGTGTAAATTAACAGAGCTCATCGAGTTCAGCGAAACATTAAGTTGACATTGAT 240
Db 769 GACAAAGTGTAAATTAACAGAGCTCATCGAGTTCAGCGAAACATTAAGTTGACATTGAT 828
Qy 241 TGCAGTCAGAGTGTGATGTTAGTAAAGTCCCAAGATGTTGGCTCTCTCTTCA 300
Db 829 TGCAGTCAGAGTGTGATGTTAGTAAAGTCCCAAGATGTTGGCTCTCTCTTCA 888
Qy 301 GACTGTTTCTCACTGTAATCTGAGTAACTGAGAGAGAGCTTCAACTGTTCACCTG 360
Db 889 GACTGTTTCTCACTGTAATCTGAGTAACTGAGAGAGAGCTTCAACTGTTCACCTG 948
Qy 361 CTTCAGAGTGGTGTGCTGAGTCTCTGCTGTTTCCCTGCTCTGTGAGAGAAATCC 420
Db 949 CTTCAGAGTGGTGTGCTGAGTCTCTGCTGTTTCCCTGCTCTGTGAGAGAAATCC 1008
Qy 421 ACTCTAGTCCCTACCTGCTATTTCTCAGCCTTGCTTACCTGTCGAACATTTGGGCCAAC 480
Db 1009 ACTCTAGTCCCTACCTGCTATTTCTCAGCCTTGCTTACCTGTCGAACATTTGGGCCAAC 1068
Qy 481 GGAATTTCTCCCAATCTTAACTTGGCTGCGAGAGATGCTCTCAACAGAGCTGATA 540
Db 1069 GGAATTTCTCCCAATCTTAACTTGGCTGCGAGAGATGCTCTCAACAGAGCTGATA 1128
Qy 541 CAGCAGAAATGGAAATGGTTAATGCTTAAATGCACTATACCTGTCCAAAGCTGACTTT 600
Db 1129 CAGCAGAAATGGAAATGGTTAATGCTTAAATGCACTATACCTGTCCAAAGCTGACTTT 1188
Qy 601 ATCCCGAGTCTGATTTCTGCTGCTGCTGCTGTAATGACAGCTTTGTGAGAAATTTTG 660
Db 1189 ATCCCGAGTCTGATTTCTGCTGCTGCTGCTGTAATGACAGCTTTGTGAGAAATTTTG 1248
Qy 661 CCGTGGTTGACAATACAGTATGATTTCAATGAGAAAGCAAAAGCTCCAAATGATGTGT 720
Db 1249 CCGTGGTTGACAATACAGTATGATTTCAATGAGAAAGCAAAAGCTCCAAATGATGTGT 1308
Qy 721 CTAGTGACATGTTTACTGAGGATCTCCGCTCCGCAACATGCTATGCTCAATCATG 780
Db 1309 CTAGTGACATGTTTACTGAGGATCTCCGCTCCGCAACATGCTATGCTCAATCATG 1368
Qy 781 AAGAGATGAGATGCTTTTGAATGAGAGCTTACAGATTTGTGAAGAAAAAGACTACT 840
Db 1369 AAGAGATGAGATGCTTTTGAATGAGAGCTTACAGATTTGTGAAGAAAAAGACTACT 1428
Qy 841 ATATCTCCAAATCTTCAATTTTCTGGCCCAACTCTGAGCTATGAGAGAAATTAAGAAC 900
Db 1429 ATATCTCCAAATCTTCAATTTTCTGGCCCAACTCTGAGCTATGAGAGAAATTAAGAAC 1488

QY 901 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAGCTGCTGACCTGGAGGAGCAAAAT 960
DB 1489 CAGACTGAGCATCAGGCGCAAGAGCAAACTCAGCTGCTGACCTGGAGGAGCAAAAT 1548
QY 961 GAACTGTCCCTGCTGCTCAGAGGGTGAACAGAAAAAGCAGAGCGCCCTCACTCCACCC 1020
DB 1549 GAACTGTCCCTGCTGCTCAGAGGGTGAACAGAAAAAGCAGAGCGCCCTCACTCCACCC 1608
QY 1021 TGTGCGCATCTGCTGCTCAGAGGAGCAGAGCAAAAGCCCTGCAATCCCGCAGCGTG 1080
DB 1609 TGTGCGCATCTGCTGCTCAGAGGAGCAGAGCAAAAGCCCTGCAATCCCGCAGCGTG 1668
QY 1081 CCCAGGCTGCGCAGCGCTGAGCGCTGTTAGAGGAGCGCCGCTGAGCAGGAGGCTC 1140
DB 1669 CCCAGGCTGCGCAGCGCTGAGCGCTGTTAGAGGAGCGCCGCTGAGCAGGAGGCTC 1728
QY 1141 AGTGGGCTGCACTGCTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCCTTCTTC 1200
DB 1729 AGTGGGCTGCACTGCTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCCTTCTTC 1788
QY 1201 TCTCTGAGATATCAATCAGTTTATATTCAGCCAGCATGCGACATCTTACATGCGCTTC 1260
DB 1789 TCTCTGAGATATCAATCAGTTTATATTCAGCCAGCATGCGACATCTTACATGCGCTTC 1848
QY 1261 TCTCTATCAGAAATGCTTTGGAAATCTAACAACCTTCCATCTCTGAGATGGGACCAAC 1320
DB 1849 TCTCTATCAGAAATGCTTTGGAAATCTAACAACCTTCCATCTCTGAGATGGGACCAAC 1908
QY 1321 AAGCTATGCGAGTTCTCCCTGTTACAGAACTATCGAGAGAGCTCCGAAACCAAGTCTCT 1380
DB 1909 AAGCTATGCGAGTTCTCCCTGTTACAGAACTATCGAGAGAGCTCCGAAACCAAGTCTCT 1968
QY 1381 GATTAAGAGAGAACCCAGCATCCCAAGAACTGACAGCCGACGCTTTCAGACAGCCAG 1440
DB 1969 GATTAAGAGAGAACCCAGCATCCCAAGAACTGACAGCCGACGCTTTCAGACAGCCAG 2028
QY 1441 AGAAACGAGATTCATTCGGTCAAGAACAGAGACAGAGAGCGCCCAAGAGTCCCTTTTA 1500
DB 2029 AGAAACGAGATTCATTCGGTCAAGAACAGAGACAGAGAGCGCCCAAGAGTCCCTTTTA 2088
QY 1501 TCTCCACTGATGAAATGAGGAGAGAGAGCAATTAACACAGCAGCTTCTTTTCGCG 1560
DB 2089 TCTCCACTGATGAAATGAGGAGAGAGAGCAATTAACACAGCAGCTTCTTTTCGCG 2148
QY 1561 CTTTCCACGAGCCAGCAGCATCTCAAGAACTGCTGAGCTGAGCTTAAAGGGCTGACAC 1620
DB 2149 CTTTCCACGAGCCAGCAGCATCTCAAGAACTGCTGAGCTTAAAGGGCTGACAC 2208
QY 1621 TCGGATATCTTTGGCCCCCAGACCTCTAACCCTTCCCTGACAGCAGCTGATTTTGGCC 1680
DB 2209 TCGGATATCTTTGGCCCCCAGACCTCTAACCCTTCCCTGACAGCAGCTGATTTTGGCC 2268
QY 1681 ACAGAGCTCTCAGACCTCTACTGCTCAGCCATCTACAGGAGGAGCTGCACTTCTCT 1740
DB 2269 ACAGAGCTCTCAGACCTCTACTGCTCAGCCATCTACAGGAGGAGCTGCACTTCTCT 2328
QY 1741 GCTTACAGCTGACAGCAGCTGCCACTTGGGAGACCAAGTCTATTTCTGTGCGCAGGCG 1800
DB 2329 GCTTACAGCTGACAGCAGCTGCCACTTGGGAGACCAAGTCTATTTCTGTGCGCAGGCG 2388
QY 1801 CAGAAAGCAGAGTACAGAGCTGACTGCGCGCGAGAGCTGGCATTAAGAGAGCCCTTTTGA 1860
DB 2389 CAGAAAGCAGAGTACAGAGCTGACTGCGCGCGAGAGCTGGCATTAAGAGAGCCCTTTTGA 2448
QY 1861 AAGCAGTTTAAAGCAGAGCTGCAAAATGAAATTTGAGAGAGCATCATGTAGAGAAC 1920
DB 2449 AAGCAGTTTAAAGCAGAGCTGCAAAATGAAATTTGAGAGAGCATCATGTAGAGAAC 2508
QY 1921 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTAGCTTTTTCGGGACAGATGAA 1980
DB 2509 AGGTCAAGGAGAGAGCTGGGAGAAAGTGGGAGTCACTAGCTTTTTCGGGACAGATGAA 2568
QY 1981 ATCATTGAGGTCTCC 1995

DB 2569 ATCATTGAGGTCTCC 2583
RESULT 23
AAH99685
ID AAH99685 standard; cDNA; 2966 BP.
AC AAH99685;
DT 16-OCT-2001 (first entry)
DE Human protein encoding cDNA sequence SEQ ID NO:520.
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
XX antiinflammatory; antirheumatic; antiallergic; immunosuppressive;
XX antibacterial; endocrine; cardiac; central nervous system; viricide;
XX anti-HIV; fungicide; antitumor; cardiovascular; antianemic; anemia;
XX antiaggregant; haemostatic; vulnery; antitumor; osteopathic; eczema;
XX dermatological; antiallergic; antianemic; antidiabetic; cytostatic;
XX neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
XX immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
XX antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
XX cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
XX genetic disease; haematopoietic disorder; platelet disorder; asthma;
XX thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
XX allergic rhinitis; diabetes; multiple sclerosis; depression;
XX Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX neurological disorder; ss.
OS Homo sapiens.
XX
XX WO200153455-A2.
XX
XX 26-JUL-2001.
XX
XX 22-DEC-2000; 2000WO-US035017.
XX
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-00488725.
XX 25-APR-2000; 2000US-00552317.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
PI WPI; 2001-457603/49.
XX
XX P-PSDB; AAM25744.
XX
XX Isolated human polynucleotides encoding polypeptides, useful for the
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
XX Claim 1; Page 578; 1217pp; English.
XX
XX AAH99166 to AAH99904 encode the human proteins given in AAM2525 to
XX AAM25963. The proteins can have activities based on the tissues and cells
XX they are expressed in, such as: antineoplastic; antirheumatic;
XX antiallergic; immunosuppressive; antibacterial; endocrine; cardiac;
XX central nervous system; viricide; anti-HIV; fungicide; antitumor;
XX cardiovascular; antianemic; antiaggregant; haemostatic; vulnery;
XX anticancer; osteopathic; dermatological; antiallergic; antianemic;
XX antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
XX antiparkinsonian; and immunostimulant. The proteins and polynucleotides
XX encoding them can be used in gene therapy, antisense therapy and vaccine
XX production. The proteins and polynucleotides are useful for screening for
XX agonists or antagonists of a protein and for the treatment and diagnosis
XX of disorders associated with the activity of a protein e.g. inflammation,
XX rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
XX neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
XX infections, autoimmunity, genetic diseases, haematopoietic disorders,
XX anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
XX osteoporosis, severe combined immunodeficiency, eczema, allergic
XX rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,

CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
XX
SQ Sequence 2966 BP; 809 A; 735 C; 688 G; 733 T; 0 U; 1 Other;

Query Match 92.7%; Score 1849; DB 4; Length 2966;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1949; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 45 GGTGGCTCTGCTGGAAAAGTGGAAAGGAAAAGTCTGCTAATTGATAGCCGGCCATTGT 104
Db 67 GGTGGCTCTGCTGGAAAAGTGGAAAGGAAAAGTCTGCTAATTGATAGCCGGCCATTGT 126
QY 105 GGAATACAAATACATCCCAATTTTGAAGCCATTATATCAACGTCGCCAAGCTTATGAA 164
Db 127 GGAATACAAATACATCCCAATTTTGAAGCCATTATATCAACGTCGCCAAGCTTATGAA 186
QY 165 GCGAAGGTGGCAACAGAACAAAGTTAATTAACAGAGCTCATCCAGCATTCAGCGAAACA 224
Db 187 GCGAAGGTGGCAACAGAACAAAGTTAATTAACAGAGCTCATCCAGCATTCAGCGAAACA 246
QY 225 TAAAGTTGACATGATTTGACAGTCAAGAGTTTATAGTTATAGATCAAAAGCTCCCAAGATGT 284
Db 247 TAAAGTTGACATGATTTGACAGTCAAGAGTTTATAGTTATAGATCAAAAGCTCCCAAGATGT 306
QY 285 TGCCTCTCTCTCTCAGACTGTTTCTCACTGACTCTTGGGTAAACTGGAGAAGACTT 344
Db 307 TGCCTCTCTCTCTCAGACTGTTTCTCACTGACTCTTGGGTAAACTGGAGAAGACTT 366
QY 345 CAACCTCTGCTACCTGCTTGCAGAGTGGTTCGTCAGATTCTCGTTGTTTCCGTGGCCT 404
Db 367 CAACCTCTGCTACCTGCTTGCAGAGTGGTTCGTCAGATTCTCGTTGTTTCCGTGGCCT 426
QY 405 CTGTGAAGGAAATTCACCTAGTCCCTACCTGACCTTTCAGCCCTTGGCTTACCTGTTCG 464
Db 427 CTGTGAAGGAAATTCACCTAGTCCCTACCTGACCTTTCAGCCCTTGGCTTACCTGTTCG 486
QY 465 CAACATTTGGCCCAACCCGAAATCTTCCCAATCTTTATCTTGGGTGCAACGAGATGTCT 524
Db 487 CAACATTTGGCCCAACCCGAAATCTTCCCAATCTTTATCTTGGGTGCAACGAGATGTCT 546
QY 525 CAACAGAGAGCTGATACAGACAAATGGGATTTGTTATGTTAATATGCAAGCTTATACCTG 584
Db 547 CAACAGAGAGCTGATACAGACAAATGGGATTTGTTATGTTAATATGCAAGCAATACCTG 606
QY 585 TCCAAAGCCTGACTTATATCCCGAGTCTCATTTCCGCGTGGCCGTGAATGAAGACTT 644
Db 607 TCCAAAGCCTGACTTATATCCCGAGTCTCATTTCCGCGTGGCCGTGAATGAAGACTT 666
QY 645 TTGTGGAATAATTTTGGCGTGGTTGACAATCAGTGAATTTCAATGAGAAAGCAAAAGC 704
Db 667 TTGTGGAATAATTTTGGCGTGGTTGACAATCAGTGAATTTCAATGAGAAAGCAAAAGC 726
QY 705 CTCGATGATGTTGTTCTAGTGCACTGTTTACCTGGGATCTCCCGCTCCGCCACATGCTG 764
Db 727 CTCGATGATGTTGTTCTAGTGCACTGTTTACCTGGGATCTCCCGCTCCGCCACATGCTG 786
QY 765 TATCGCTCATCATCATGAGAGAGTGAACATGCTTTAATGATGAAGCTTACAGATTTGTGAA 824
Db 787 TATCGCTCATCATCATGAGAGAGTGAACATGCTTTAATGATGAAGCTTACAGATTTGTGAA 846
QY 825 AGAAAAAGAGCTTACTATATCTCCAAACTTCAATTTTCTGGCCCAACTCTGGAATATGA 884
Db 847 AGAAAAAGAGCTTACTATATCTCCAAACTTCAATTTTCTGGCCCAACTCTGGAATATGA 906
QY 885 GAAAGAAGATTGAAGCAAGCTGAGCAATCAAGGGCCAAAGAGCAAACTCAAGCTGTGCA 944
Db 907 GAAAGAAGATTGAAGCAAGCTGAGCAATCAAGGGCCAAAGAGCAAACTCAAGCTGTGCA 966
QY 945 CCTGAGAGAGCCAAATGAACCTGTCCCTGCTGCTCAGAGGGTGAAGAGAAAGCGAGAC 1004
Db 967 CCTGAGAGAGCCAAATGAACCTGTCCCTGCTGCTCAGAGGGTGAAGAGAAAGCGAGAC 1026
```

```
QY 1005 GCCCCTCACTCCACCTGTGCGGACTGTGTAACCTCAAGAGGACAGAGCAAAAGCCCGT 1064
Db 1027 GCCCCTCACTCCACCTGTGCGGACTGTGTAACCTCAAGAGGACAGAGCAAAAGCCCGT 1086
QY 1065 GCATCCCGCAGAGGCGCCAGGCTGCGCCAGGCTGAGAGCCGTGCTTTAGAGGACAGCCC 1124
Db 1087 GCATCCCGCAGAGGCGCCAGGCTGCGCCAGGCTGAGAGCCGTGCTTTAGAGGACAGCCC 1146
QY 1125 GCTGATACAGGCGCTCAGTGGCTGCACTGTGCCAGACAGAGCTGGAAGACAGCAATTA 1184
Db 1147 GCTGATACAGGCGCTCAGTGGCTGCACTGTGCCAGACAGAGCTGGAAGACAGCAATTA 1206
QY 1185 GCTCAAGGCTTCTTCTCTCTGAGATTAATCACTTCAATTTCAATTCAGCCAGCATGCGACG 1244
Db 1207 GCTCAAGGCTTCTTCTCTGAGATTAATCACTTCAATTTCAATTCAGCCAGCATGCGACG 1266
QY 1245 ATCTTAATAGGCTTCTCTCATCAGAGAGTGTGGAAATATCAAAACCTTCACATAC 1304
Db 1267 ATCTTAATAGGCTTCTCTCATCAGAGAGTGTGGAAATATCAAAACCTTCACATAC 1326
QY 1305 TCTGATGGAACCAACAAAGCTATGCAAGTTCTCCCTGTTTCAGGAACATATGAGACAGAC 1364
Db 1327 TCTGATGGAACCAACAAAGCTATGCAAGTTCTCCCTGTTTCAGGAACATATGAGACAGAC 1386
QY 1365 TCCGAAACCAAGTCTGTAAGAGAGAAAGCCAGATCCCAAGAGCTGCAGACCGCCAG 1424
Db 1387 TCCGAAACCAAGTCTGTAAGAGAGAAAGCCAGATCCCAAGAGCTGCAGACCGCCAG 1446
QY 1425 GCCTTCAGACAGCCAGACCAAGCATTTGCAATTCGCTCAGAACCCAGACAGTGGCACCGC 1484
Db 1447 GCCTTCAGACAGCCAGACCAAGCATTTGCAATTCGCTCAGAACCCAGACAGTGGCACCGC 1506
QY 1485 CCAAGATGCTCTTTTATCTCACTGCATCGAATGGAGCGGTGAGGACAAATTAACAAC 1544
Db 1507 CCAAGATGCTCTTTTATCTCACTGCATCGAATGGAGCGGTGAGGACAAATTAACAAC 1566
QY 1545 CAGCTTCCTTTTGGCCCTTTCACACAGCCAGACACTTACAGAAAGTCTGTGGCTGGG 1604
Db 1567 CAGCTTCCTTTTGGCCCTTTCACACAGCCAGACACTTACAGAAAGTCTGTGGCTGGG 1626
QY 1605 CTTTAAGGGCTGGGCACTCGGATATCTTTGGCCCGCCAGACCTTAACCCCTCCCTGAC 1664
Db 1627 CTTTAAGGGCTGGGCACTCGGATATCTTTGGCCCGCCAGACCTTAACCCCTCCCTGAC 1686
QY 1665 CAGTGGTATTTTGGCCAGAGATCTCAACATTTCTATCTGTCTCAGCCATCTTACGAGG 1724
Db 1687 CAGTGGTATTTTGGCCAGAGATCTCAACATTTCTATCTGTCTCAGCCATCTTACGAGG 1746
QY 1725 CAGTGCAGTTAATCTGCTTACAGCTGACGCGAGCTGCCCACTTGGGGAGACCAAGTTTA 1784
Db 1747 CAGTGCAGTTAATCTGCTTACAGCTGACGCGAGCTGCCCACTTGGGGAGACCAAGTTTA 1806
QY 1785 TTCTGTGCGGAGCGGCGAGAGCCAAAGTGAACAGCTGACTCGCGCGAGAGCTGACATGA 1844
Db 1807 TTCTGTGCGGAGCGGCGAGAGCCAAAGTGAACAGCTGACTCGCGCGAGAGCTGACATGA 1866
QY 1845 AGAGAGCCCTTTGAAAAGAGATTAAAGCAGAACCTGCCAAATGGAATTTTGAAGAGAG 1904
Db 1867 AGAGAGCCCTTTGAAAAGAGATTAAAGCAGAACCTGCCAAATGGAATTTTGAAGAGAG 1926
QY 1905 CATCATGTCAAGAAACAGTCAACGGGAAGAGCTGGGGAAAGTGGGCAAGTCACTTACGCTT 1964
Db 1927 CATCATGTCAAGAAACAGTCAACGGGAAGAGCTGGGGAAAGTGGGCAAGTCACTTACGCTT 1986
QY 1965 TTGGGCGACATGGAATCATTTGAGGTCTCC 1995
Db 1987 TTGGGCGACATGGAATCATTTGAGGTCTCC 2017
```

RESULT 24
AABD9492
ID AABD9492 standard; DNA; 2732 BP.
XX

AC AAD09492;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human SGP002 phosphatase polypeptide encoding DNA.
XX
KM Human: SGP002 phosphatase polypeptide; phosphatase-related disease;
KM infection-related disorder; ocular disease; organ transplant rejection;
KM infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KM metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
KM Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KM cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KM attention disorder; cognitive disorder; psychotic disorder; cytostatic;
KM neurological disorder; viricide; nocotropic; cerebroprotective; therapy;
KM neuroprotective; antibacterial; antiviral; tranquilliser; antischismatic;
KM hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KM antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KM MKP; migraline; chromosome 12p11.1-p12.1; ds.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 538..2535
FT /*tag= a
FT /product= "Human SGP002 phosphatase polypeptide"
XX
XX MO200146394-A2.
XX
PD 28-JUN-2001.
XX
PF 21-DEC-2000; 2000MO-US034736.
XX
PR 21-DEC-1999; 99US-0173255P.
PR 28-DEC-1999; 99US-0175766P.
PR 25-JAN-2000; 2000US-0178078P.
PR 31-JAN-2000; 2000US-0179301P.
XX
XX (SUGEN-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S;
XX Hill RJ, Piansagan P;
XX
XX MPI: 2001-418058/44.
XX P-PSDB; AAE04834.
XX
XX Novel phosphatase polypeptide useful for treating cancers, immune-related
XX diseases and disorder, cardiovascular disease, brain or neuronal-
XX associated diseases and metabolic disorders.
XX
XX Claim 29; Fig 1; 186pp; English.
XX
XX The present invention relates to phosphatase polypeptides, nucleotide
XX sequences encoding them, as well as various products and methods useful
XX for the diagnosis and treatment of various phosphatase-related diseases
XX and conditions. Substance that modulates the activity of phosphatase
XX polypeptide is used to treat immune-related diseases and disorders,
XX cardiovascular disease, brain or neuronal-associated diseases and
XX metabolic disorders, including cancers of tissues, cancers of
XX haematopoietic origin, diseases of central and peripheral nervous system,
XX Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
XX lateral sclerosis, viral infections, infections caused by prions,
XX bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual
XX dysfunction, mood disorders, attention disorders, cognitive disorders,
XX hypotension, hypertension, psychotic disorders, neurological disorders,
XX dyskinesias and organ transplant rejection. The present sequence is a DNA
XX encoding human SGP002 phosphatase polypeptide. This sequence is
XX classified as dual specificity phosphatase (DSP) and MAP kinase
XX phosphatase (MKP). SGP002 gene maps to chromosome1 position 12p11.1-p12.1
XX
SQ Sequence 2732 BP; 710 A; 684 C; 686 G; 652 T; 0 U; 0 Other;

Query Match 92.3%; Score 1842; DB 4; Length 2732;
Best Local Similarity 99.8%; Pred. No. 0;

Matches 1992; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATTGAACTCAAAATTGTTACTGAGAGTTGGTGGCTCTGCTGGAA 60
DB 538 ATGGCCCATGAGATGATTGAACTCAAAATTGTTACTGAGAGTTGGTGGCTCTGCTGGAA 597
QY 61 AGTGAAGGAAAAAGTGTGCTGAATGATAGCCGGCCATTTGTGATATACATACATAC 120
DB 598 AGTGAAGGAAAAAGTGTGCTGAATGATAGCCGGCCATTTGTGATATACATACATAC 657
QY 121 CACATTTTGAAGCCATTATATCAACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 180
DB 658 CACATTTTGAAGCCATTATATCAACTGCTCCAGCTTATGAGCGAAGTTGCAACAG 717
QY 181 GACAAAGTTTAATTACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATTAT 240
DB 718 GACAAAGTTTAATTACAGAGCTCATCCAGCATTCAGCGAACAATAAGTTGACATTAT 777
QY 241 TGCAGTCAGAGGTTGATTTAACAATCAAGCTCCAGAGTTGGCTCTGCTCTTCA 300
DB 778 TGCAGTCAGAGGTTGATTTAACAATCAAGCTCCAGAGTTGGCTCTGCTCTTCA 837
QY 301 GACTGTTTCTCACTGTAATCTTGGGTAACAGAGAGCTTCAACTCTGTTCACTG 360
DB 838 GACTGTTTCTCACTGTAATCTTGGGTAACAGAGAGCTTCAACTCTGTTCACTG 897
QY 361 CTTGCAAGTGGGTTTGTGAGATTCTCTGCTGTTTCCCTGGCCTCTGTAAGAAATCC 420
DB 898 CTTGCAAGTGGGTTTGTGAGATTCTCTGCTGTTTCCCTGGCCTCTGTAAGAAATCC 957
QY 421 ACTCTAGTCCCTACCTGATTTCTCAAGCTTCTACCTGTTGCCAACAATTGGGCCAAC 480
DB 958 ACTCTAGTCCCTACCTGATTTCTCAAGCTTCTACCTGTTGCCAACAATTGGGCCAAC 1017
QY 481 CCAATTCCTCCCAATCTTATCTTGGCTCCAGAGAGTGTCTTCAACAAGAGCTGATG 540
DB 1018 CCAATTCCTCCCAATCTTATCTTGGCTCCAGAGAGTGTCTTCAACAAGAGCTGATG 1077
QY 541 CAGCAAGATGGAGTTGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCCTGACTT 600
DB 1078 CAGCAAGATGGAGTTGTTATGTTAAATGCGAGCTATACCTGTCCAAAGCCTGACTT 1137
QY 601 ATCCCGAGATCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 1138 ATCCCGAGATCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
QY 661 CCGTGGTTGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTT 720
DB 1198 CCGTGGTTGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCAAATGATGTT 1257
QY 721 CTATGTCATCTTTAGCTGGATCTCCGCTCCGCCACCATGCTATGCTTACATCATG 780
DB 1258 CTATGTCATCTTTAGCTGGATCTCCGCTCCGCCACCATGCTATGCTTACATCATG 1317
QY 781 AAGAGATGACATGCTTTAATGATGAGCTTACAGTTTGTGAAAAAGAAAGAAAGAAAG 840
DB 1318 AAGAGATGACATGCTTTAATGATGAGCTTACAGTTTGTGAAAAAGAAAGAAAGAAAG 1377
QY 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGAAAGAAAG 900
DB 1378 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGAAAGAAAG 1437
QY 901 CAGACTGAGCATCAGAGGCCAAAGCAAACTCAAGCTGCTGACCTGAGAGAGCAAAAT 960
DB 1438 CAGACTGAGCATCAGAGGCCAAAGCAAACTCAAGCTGCTGACCTGAGAGAGCAAAAT 1497
QY 961 GAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 1498 GAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1557
QY 1021 TGTGCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1558 TGTGCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1617

QY	1081	CCGACGCTGCCACGCTGCGACGCGCTGCTGTTAGAGAGACGCCGCTGTATCAGACGCTC	1140
Db	1618	CCGACGCTGCCACGCGCGACGCGCTGCTGTTAGAGAGACGCCGCTGTATCAGACGCTC	1677
QY	1141	AGTGGCGTGCACCTGTTCGCGACGACGACTGGAGAACAGCAATTAACCTCAAGCTTCCCTTC	1200
Db	1678	AGTGGCGTGCACCTGTTCGCGACGACGACTGGAGAACAGCAATTAACCTCAAGGCTTCCCTTC	1737
QY	1201	TCCTCGATATCAATATCAGTTTCATATTCATTCAGCCAGCATGGCAGCATCTTTCATATGCTTC	1260
Db	1738	TCCTCGATATCAATATCAGTTTCATATTCATTCAGCCAGCATGGCAGCATCTTTCATATGCTTC	1797
QY	1261	TCCTCATTCAGAAAGATCTTTGGAAATCTTAACAAACCTTCACATCTCTTGATGGAGCCAC	1320
Db	1798	TCCTCATTCAGAAAGATCTTTGGAAATCTTAACAAACCTTCACATCTCTTGATGGAGCCAC	1857
QY	1321	AACTATTCGCAGTTCTCCCTGTTTCAGAAACTATCGAGCAGACTCCGGAACAGACTCT	1380
Db	1858	AACTATTCGCAGTTCTCCCTGTTTCAGAAACTATCGAGCAGACTCCGGAACAGACTCT	1917
QY	1381	GATPAGAGAAAGCCAGCATCCCGAAGAACTGCGACATCCGCGAGCGCTTCAGACAGCCAG	1440
Db	1918	GATPAGAGAAAGCCAGCATCCCGAAGAACTGCGACATCCGCGAGCGCTTCAGACAGCCAG	1977
QY	1441	AGCAAGCGATTGCATTTGGTTCAGAACACGACAGCAGCTGGCACCGCCGAGAGTCCCTTTTA	1500
Db	1978	AGCAAGCGATTGCATTTGGTTCAGAACACGACAGCAGCTGGCACCGCCGAGAGTCCCTTTTA	2037
QY	1501	TCGCCACCTGCATTCGAAAGTGGAGCGCTGGAGAGCAATTAACACACACAGCTTCTTTTCGCG	1560
Db	2038	TCGCCACCTGCATTCGAAAGTGGAGCGCTGGAGAGCAATTAACACACACAGCTTCTTTTCGCG	2097
QY	1561	CTTTCACCCAGCCAGCAGACACTTCACGAAATCTGTCTGAGCTGGGCTTAAAGGCTGGCAC	1620
Db	2098	CTTTCACCCAGCCAGCAGACACTTCACGAAATCTGTCTGAGCTGGGCTTAAAGGCTGGCAC	2157
QY	1621	TCGGAATTCCTGGCCCCCGACAACTCTTAACCCCTTCCCTTAACACAGCACTGTAATTTTGGC	1680
Db	2158	TCGGAATTCCTGGCCCCCGACAACTCTTAACCCCTTCCCTTAACACAGCACTGTAATTTTGGC	2217
QY	1681	ACAGAGTCTTCACATTTCTACTCTGTGCTCAGCCACTTCACGAGAGCAGTGCAGATTACTCT	1740
Db	2218	ACAGAGTCTTCACATTTCTACTCTGTGCTCAGCCACTTCACGAGAGCAGTGCAGATTACTCT	2277
QY	1741	GCCTACACCTGCACGCTGAGTCCCACTTGGAGAACCAATGTAATTTCTGTGCGCAGGCGG	1800
Db	2278	GCCTACACCTGCACGCTGAGTCCCACTTGGAGAACCAATGTAATTTCTGTGCGCAGGCGG	2337
QY	1801	CAGAAGCCAATGTACAGAGCTGACTGCGCGCGGAGCTGGCATTAAGAGAGCCCTTTTGA	1860
Db	2338	CAGAAGCCAATGTACAGAGCTGACTGCGCGCGGAGCTGGCATTAAGAGAGCCCTTTTGA	2397
QY	1861	AAGCAATTTTAAAGCGACAGAGCTGCAGAAATTTGGAGAGAGCATATGTCAAGAAAC	1920
Db	2398	AAGCAATTTTAAAGCGACAGAGCTGCAGAAATTTGGAGAGAGCATATGTCAAGAAAC	2457
QY	1921	AGGTCAAGGGAAGAGCTGGGGAAAGTGGGAGAGCATCTTAAGCTTTTCGGGCGACATGAA	1980
Db	2458	AGGTCAAGGGAAGAGCTGGGGAAAGTGGGAGAGCATCTTAAGCTTTTCGGGCGACATGAA	2517
QY	1981	ATCATTTAGGCTTCC 1995	
Db	2518	ATCATTTAGGCTTCC 2532	

RESULT 25
ABN59704
ID ABN59704 strand; cDNA; 3104 BP
XX
XX ABN59704;
XX
DT 28-JUN-2002 (first entry)

XX	Novel human coding sequence SEQ ID NO: 115.
DE	Human, anti-nausea; vulnery; anti-inflammatory; immunomodulatory;
KW	anti-infectivity; cerebroprotective; cytosolic; rheumatic; gene therapy;
KM	neuroprotective; antiparkinsonian; protein therapy; EST;
KW	expressed sequence tag; gene, ss.
XX	
OS	Homo sapiens.
XX	
XX	WO200222660-A2.
PD	21-MAR-2002.
XX	
XX	10-SEP-2001; 2001WO-US026015.
PF	
XX	11-SEP-2000; 2000US-00659671.
PR	
XX	(HYSE-) HYSE INC.
PA	
PI	Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI	Xue AJ, Yang Y, Wehrman T, Dymac RT;
XX	
DR	WPI: 2002-292408/33.
XX	P-PSDB; ABB97291.
PT	An isolated polynucleotide for treating diseases associated with its
XX	encoded polypeptide such as cancer and multiple sclerosis.
XX	
PS	Claim 1; SEQ ID NO 115; 509pp; English.
XX	
CC	The present invention provides the protein and coding sequences of 444
CC	novel human proteins. These were isolated from expressed sequences tags
CC	(ESTs). They can be used to stimulate cell growth, to regulate
CC	haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC	e.g. in burn treatment, to regulate the immune system e.g. to treat
CC	multiple sclerosis, to regulate activin or inhibit e.g. to treat
CC	infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC	and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC	rheumatoid arthritis, and to treat nervous system disorders e.g.
CC	Parkinson's disease. The present sequence is a coding sequence of the
CC	invention
XX	
XX	Sequence 3104 BP; 845 A; 772 C; 718 G; 769 T; 0 U; 0 Other;
SO	
Query Match	92.3%; Score 1842; DB 6; Length 3104;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 1992; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	1 ATGGCCCATGAGATGATTTGAACCTCAATTGTACTGAGAGTTGGTGCCTGCGAA 60
DB	159 ATGGCCCATGAGATGATTTGAACCTCAATTGTACTGAGAGTTGGTGCCTGCGAA 218
QY	61 AGTGAACGGAAAAAGCTGCTTAATTGATAGCGGCGCATTTGTGAATATACATATACC 120
DB	219 AGTGAACGGAAAAAGCTGCTTAATTGATAGCGGCGCATTTGTGAATATACATATACC 278
QY	121 CACATTTTGAAGCATTATATATCACTGCTCCAAGCTTATGAAGCGAAGTTGCCAACG 180
DB	279 CACATTTTGAAGCATTATATATCACTGCTCCAAGCTTATGAAGCGAAGTTGCCAACG 338
QY	181 GACAAAGTGTTAATTAACAGAGCTATCAGCATTTACGCGAAACATTAAGTTGACATTGAT 240
DB	339 GACAAAGTGTTAATTAACAGAGCTATCAGCATTTACGCGAAACATTAAGTTGACATTGAT 398
QY	241 TGCAGTCGAAAGTTGATTTAGATTCAGATCAAGCTCCCAATGTTGCTCTCTCTTCA 300
DB	399 TGCAGTCGAAAGTTGATTTAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 458
QY	301 GACGTTTCTCACTGATCTTCTGGGTTAACTGAGAGAGCTTCACTCTGTTCACTTG 360
DB	459 GACGTTTCTCACTGATCTTCTGGGTTAACTGAGAGAGCTTCACTCTGTTCACTTG 518

QY 361 CTTCAGAGTGGGTTTGGTGAAGTCTCTGTTGTTTCCCTGAGCTCTGTAAGGAAATCC 420
 DB 519 CTTCAGAGTGGGTTTGGTGAAGTCTCTGTTGTTTCCCTGAGCTCTGTAAGGAAATCC 578
 QY 421 ACTTACTGCTTACCTGCAATTTCTCAGCCTTGTCTTACCTGTGCAACATTTGGCCACC 480
 DB 579 ACTTACTGCTTACCTGCAATTTCTCAGCCTTGTCTTACCTGTGCAACATTTGGCCACC 638
 QY 481 CGAATTTCTTCCCAATCTTTATCTTTGGTGGCAGAGAGATGTCTCTCAACAGAGCTGATA 540
 DB 639 CGAATTTCTTCCCAATCTTTATCTTTGGTGGCAGAGAGATGTCTCTCAACAGAGCTGATA 698
 QY 541 CAGAGAAATGGAGATGGTATGTTGTTAAATGCCAGCTATACCTGTCCAAAGCCGACTTT 600
 DB 699 CAGAGAAATGGAGATGGTATGTTGTTAAATGCCAGCTATACCTGTCCAAAGCCGACTTT 758
 QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAAATTTTG 660
 DB 759 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAAATTTTG 818
 QY 661 CCGTGTGTGCAAAATAGTAGATTTTATGAGAAAGCAAAAGCTTCAATGATGTT 720
 DB 819 CCGTGTGTGCAAAATAGTAGATTTTATGAGAAAGCAAAAGCTTCAATGATGTT 878
 QY 721 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCCACATGCTATGCTCATCATG 780
 DB 879 CTAGTCACTGTTTAACTGGGATCTCCGCTCCGCCACATGCTATGCTCATCATG 938
 QY 781 AAGAGATGACATGTTTATGAGAGCTTAAAGATTTGTGAAGAAAGAAAGCACTACT 840
 DB 939 AAGAGATGACATGTTTATGAGAGCTTAAAGATTTGTGAAGAAAGAAAGCACTACT 998
 QY 841 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGTGACTATGAGAAAGATTAAGAAC 900
 DB 999 ATATCTCAAACTTCAATTTTCTGGGCAACTCTGTGACTATGAGAAAGATTAAGAAC 1058
 QY 901 CAGACTGGAGCATCAGGGCCAAAGAGCAACTCAAGCTGTGACCTGGAGAGCCAAAT 960
 DB 1059 CAGACTGGAGCATCAGGGCCAAAGAGCAACTCAAGCTGTGACCTGGAGAGCCAAAT 1118
 QY 961 GAACTGTCTCTGCTCTCAGAGGTGACAGAAAGCGCCCTCAGTCCACC 1020
 DB 1119 GAACTGTCTCTGCTCTCAGAGGTGACAGAAAGCGCCCTCAGTCCACC 1178
 QY 1021 TGTGCGACTCTGCTCTCAGAGGTGACAGAAAGCGCCCTCAGTCCACC 1080
 DB 1179 TGTGCGACTCTGCTCTCAGAGGTGACAGAAAGCGCCCTCAGTCCACC 1238
 QY 1081 CCCAGGCTGCGCAGCGTGCAGCGCTGTGAGAGAGCGCGCTGTGACAGGGCTC 1140
 DB 1239 CCCAGGCTGCGCAGCGTGCAGCGCTGTGAGAGAGCGCGCTGTGACAGGGCTC 1298
 QY 1141 AGTGGGTGCACTGTCTCCGACAGAGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
 DB 1299 AGTGGGTGCACTGTCTCCGACAGAGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1358
 QY 1201 TCTCTGATATCAATATGTTTATATTTACGACAGATGAGAGATCTTCAATGCTTC 1260
 DB 1359 TCTCTGATATCAATATGTTTATATTTACGACAGATGAGAGATCTTCAATGCTTC 1418
 QY 1261 TCTCTATCAGAAATGTTTGAATCTAATACTTCCCTCTCTGATGGGACCAAC 1320
 DB 1419 TCTCTATCAGAAATGTTTGAATCTAATACTTCCCTCTCTGATGGGACCAAC 1478
 QY 1321 AAGCTATGCAATTTCTCCCTGTTCAAGAACTATCGAGAGACTCCGAAACCACTGCT 1380
 DB 1479 AAGCTATGCAATTTCTCCCTGTTCAAGAACTATCGAGAGACTCCGAAACCACTGCT 1538
 QY 1381 GATTAAGAGAGAGCAGCATCTCCCAAGAGAGCTGACAGCCGAGCTTCAAGAGCCAG 1440
 DB 1539 GATTAAGAGAGAGCAGCATCTCCCAAGAGAGCTGACAGCCGAGCTTCAAGAGCCAG 1598
 QY 1441 AGCAGGATTTGATTCGTTCAAAACAGACAGAGTGGCCAGCCGCAAGTCTTCTTTA 1500

DB 1599 AGCAGGATTTGATTCGTTCAAGAACAGAGAGTGGACAGCCGCCAGAGTCCCTTTA 1658
 QY 1501 TCTCAGCTGATGAAATGGAGCGTGAAGACATTTACCAACCAAGCTCTTTTGGC 1560
 DB 1659 TCTCAGCTGATGAAATGGAGCGTGAAGACATTTACCAACCAAGCTCTTTTGGC 1718
 QY 1561 CTTTCCAGCAGAGAGACCTCAGAGAGTCTGCTGGCTTGAAGGCTGGCAC 1620
 DB 1719 CTTTCCAGCAGAGAGACCTCAGAGAGTCTGCTGGCTTGAAGGCTGGCAC 1778
 QY 1621 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGAGCTGTAATTTTGC 1680
 DB 1779 TCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGAGCTGTAATTTTGC 1838
 QY 1681 ACAGAGTCTCAGACTTCTACTCTGCTCAGAGCATTTACGAGAGCAGTGCAGTACTCT 1740
 DB 1839 ACAGAGTCTCAGACTTCTACTCTGCTCAGAGCATTTACGAGAGCAGTGCAGTACTCT 1898
 QY 1741 GCCTAAGCTGACAGCCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGGCAGGCG 1800
 DB 1899 GCCTAAGCTGACAGCCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGGCAGGCG 1958
 QY 1801 CAGAGCCAAAGTGAAGAGCTGACTGCGCGGAGCTGACATGAAGAGAGCCCTTTGAA 1860
 DB 1959 CAGAGCCAAAGTGAAGAGCTGACTGCGCGGAGCTGACATGAAGAGAGCCCTTTGAA 2018
 QY 1861 AAGCAGTTTAAAGCAGAAAGCTGCCAATGAAATTTGAGAGAGCATATGTCAGAGAAC 1920
 DB 2019 AAGCAGTTTAAAGCAGAAAGCTGCCAATGAAATTTGAGAGAGCATATGTCAGAGAAC 2078
 QY 1921 AGGTACGGGAAAGAGTGGGAAAGTGGGAGTCACTTACCTTTTGGGACAGATGAA 1980
 DB 2079 AGGTACGGGAAAGAGTGGGAAAGTGGGAGTCACTTACCTTTTGGGACAGATGAA 2138
 QY 1981 ATCATTGAGTCTCC 1995
 DB 2139 ATCATTGAGTCTCC 2153

RESULT 26
 ADAS3105
 ID ADAS3105 standard; cDNA; 2102 BP.
 XX
 AC ADAS3105;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human coding sequence, SEQ ID 673.
 XX
 XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KW Inflammatory disease; osteoporosis; neurological disease; gene; ss.
 OS Homo sapiens.
 PN EP1293569-A2.
 PD 19-MAR-2003.
 XX
 PF 21-MAR-2002; 2002EP-00006586.
 XX
 PR 14-SEP-2001; 2001JP-00328381.
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 PA (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 PI Teogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y,
 XX
 DR WPI; 2003-395539/38.

DR P-PSDB; ADA54744.

XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 1; SEQ ID NO 673; 205pp; English.

XX The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 2102 BP; 542 A; 546 C; 510 G; 504 T; 0 U; 0 Other;

Query Match 89.8%; Score 1791; DB 7; Length 2102;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGCCATGATGATGATGAACCTCAATTGTACTGAGAGGTGGTGGCTCTCTGGAA 60
Db 56 ATGGCCATGATGATGATGAACCTCAATTGTACTGAGAGGTGGTGGCTCTCTGGAA 115
QY 61 AGTGAACGGAAAAAGTCTGCTAATTGATAGCCGCAATTGTGAAATACATACATCC 120
Db 116 AGTGAACGGAAAAAGTCTGCTAATTGATAGCCGCAATTGTGAAATACATACATCC 175
QY 121 CACATTTTGGAAAGCATTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACAG 180
Db 176 CACATTTTGGAAAGCATTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCAACAG 235
QY 181 GACAAAGTGTAAATTAACAGAGCTCATCAGCATTCAGCGAAACATTAAGTTGATTCAT 240
Db 236 GACAAAGTGTAAATTAACAGAGCTCATCAGCATTCAGCGAAACATTAAGTTGATTCAT 295
QY 241 TGCAGTGAAGAGTGTAGTTACGATCAAGCTCCCAAGATGTTCCCTCTCTCTTCA 300
Db 296 TGCAGTGAAGAGTGTAGTTACGATCAAGCTCCCAAGATGTTCCCTCTCTCTTCA 355
QY 301 GACTGTTTCTCACTGACTTCTGGGTAACTGGAAAGAGCTTCAACTCTGTCACTG 360
Db 356 GACTGTTTCTCACTGACTTCTGGGTAACTGGAAAGAGCTTCAACTCTGTCACTG 415
QY 361 CTTCAGGTGGTGTGCTGAGTTCCTGTTTTCCTGGCTCTGTGAAAGAAATCC 420
Db 416 CTTCAGGTGGTGTGCTGAGTTCCTGTTTTCCTGGCTCTGTGAAAGAAATCC 475
QY 421 ACTCTAGTCCCTACCTGCAATTTCTGAGCTTGAACCTGTTGCCAATTTGGCCAAAC 480
Db 476 ACTCTAGTCCCTACCTGCAATTTCTGAGCTTGAACCTGTTGCCAATTTGGCCAAAC 535
QY 481 CGAATTTCTCCCAATCTTATCTTGGCTCCAGCGAGATGCTCTCAAGAGCTGATA 540
Db 536 CTAAATCTTCCCAATCTTATCTTGGCTCCAGCGAGATGCTCTCAAGAGCTGATA 595
QY 541 CAGCAAAATGGAAATGGTAAATGTGTTAAATGCGAGTATACCTGTCCAAAGCTGACTT 600
Db 596 CAGCAAAATGGAAATGGTAAATGTGTTAAATGCGAGTATACCTGTCCAAAGCTGACTT 655
QY 601 ATCCCGAGTCTCAATTTCTGCGTGGCTGGAATGACGTTTGTGAAATAATTTTG 660
Db 656 ATCCCGAGTCTCAATTTCTGCGTGGCTGGAATGACGTTTGTGAAATAATTTTG 715
QY 661 CCGTGTGGCAAAATCAGTAGATTTCATGAAAGCAAAAGCTCCCAATGATGTGT 720
Db 716 CCGTGTGGCAAAATCAGTAGATTTCATGAAAGCAAAAGCTCCCAATGATGTGT 775
QY 721 CTAGTGAATGTTTGGTGGGATCTCCCGCTCCGCAACATGCTATGCTTACATATG 780
Db 776 CTAGTGAATGTTTGGTGGGATCTCCCGCTCCGCAACATGCTATGCTTACATATG 835
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAAGCTACT 840

Db 836 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAAGCTACT 895
QY 841 ATATCTCCCAATCTCAATTTTCTGGCCCACTCTGCACTATGAGAAGATTAAAGAC 900
Db 896 ATATCTCCCAATCTCAATTTTCTGGCCCACTCTGCACTATGAGAAGATTAAAGAC 955
QY 901 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTTGAAGAGCAAT 960
Db 956 CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTTGAAGAGCAAT 1015
QY 961 GAACTGTCTCTGCTGTCTCAGAGGGTGAAGAAAGCAGAGCGCCCTGATCAACCC 1020
Db 1016 GAACTGTCTCTGCTGTCTCAGAGGGTGAAGAAAGCAGAGCGCCCTGATCAACCC 1075
QY 1021 TGTGCGAATCTGTACCTCAGAGGCAAGAGCAAGAGCCCGTGCATCCGCGAGGTG 1080
Db 1076 TGTGCGAATCTGTACCTCAGAGGCAAGAGCAAGAGCCCGTGCATCCGCGAGGTG 1135
QY 1081 CCCAGCGTCCAGCGTCAAGCGCTGCTGTTAAGAGACAGCCCGTGTACAGGCGTC 1140
Db 1136 CCCAGCGTCCAGCGTCAAGCGCTGCTGTTAAGAGACAGCCCGTGTACAGGCGTC 1195
QY 1141 AGTGGCGTCACTGTCTGCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCTTCTTC 1200
Db 1196 AGTGGCGTCACTGTCTGCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCTTCTTC 1255
QY 1201 TCTGTGATTAATCAATCAGTTTCAATTTCAAGCCAGATGAGAGATCTTCAATGCTTC 1260
Db 1256 TCTGTGATTAATCAATCAGTTTCAATTTCAAGCCAGATGAGAGATCTTCAATGCTTC 1315
QY 1261 TCTCATCAGAAAGTCTTTGGAATACCAAACTTTCATCTGTGATGAGAGCAAC 1320
Db 1316 TCTCATCAGAAAGTCTTTGGAATACCAAACTTTCATCTGTGATGAGAGCAAC 1375
QY 1321 AAGCTATGCCAGTCTCCCGTTTGAAGAACTTTCGAGAGACATCCGAAACCAATGCT 1380
Db 1376 AAGCTATGCCAGTCTCCCGTTTGAAGAACTTTCGAGAGACATCCGAAACCAATGCT 1435
QY 1381 GATTAAGAGAGAGCAGACATCCCAAGAGCTGACAGCCGCTTCAAGAGCCAG 1440
Db 1436 GATTAAGAGAGAGCAGACATCCCAAGAGCTGACAGCCGCTTCAAGAGCCAG 1495
QY 1441 AGCAAGGATTTGATTTGCTGCTGCAAGACAGAGGAGTCCGCTTCAAGAGGCTT 1500
Db 1496 AGCAAGGATTTGATTTGCTGCTGCAAGACAGAGGAGTCCGCTTCAAGAGGCTT 1555
QY 1501 TCTCACTGATGAGAGTGGAGCTGAGAGCAATTTACACACAGCTTCTTTTCCGC 1560
Db 1556 TCTCACTGATGAGAGTGGAGCTGAGAGCAATTTACACACAGCTTCTTTTCCGC 1615
QY 1561 CTTTCCACAGCAGCAGACCTCAGAAATGCTGTGAGCTTGAAGGCTTGGAC 1620
Db 1616 CTTTCCACAGCAGCAGACCTCAGAAATGCTGTGAGCTTGAAGGCTTGGAC 1675
QY 1621 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCGACCAAGAGCTGTGATTTTGGC 1680
Db 1676 TCGGATATCTTGGCCCCCAGACCTTACCCCTTCCGACCAAGAGCTGTGATTTTGGC 1735
QY 1681 ACAGAGTCTCAACATTTACTCTGCTCAGCAGCATCTACAGAGGAGGAGGATTA 1740
Db 1736 ACAGAGTCTCAACATTTACTCTGCTCAGCAGCATCTACAGAGGAGGAGGATTA 1795
QY 1741 GCTTACAGTGAAGCAGCTGCGCCACTTGGGAGACCAAGTCTTGTGCGAGGCGG 1800
Db 1796 GCTTACAGTGAAGCAGCTGCGCCACTTGGGAGACCAAGTCTTGTGCGAGGCGG 1855
QY 1801 CAGAAAGCAAGTGAAGAGCTGATCGGCGGAGGAGCTGGCAATGAAGAGCCCTTTGAA 1860
Db 1856 CAGAAAGCAAGTGAAGAGCTGATCGGCGGAGGAGCTGGCAATGAAGAGCCCTTTGAA 1915
QY 1861 AAGCACTTAAACGAGAGGCTGCCAAATGGAATTTGAGAGAGCATATGTCAGAAAC 1920

Db 1916 AACGAGTTTAAACGACAGAGCTGCCAAATGGAAATTGGAGAGAGCATCATGTCAGAGAAC 1975
QY 1921 AGGTACACGGAAAGAGCTGGGGAAGTGGGACATGCTTACGTTTTCGGGACACATGGA 1980
Db 1976 AGGTACACGGGAAGAGCTGGGGAAGTGGGACATGCTTTCGGGACACATGGA 2035
QY 1981 ATCATTTAGAGTCTCC 1995
Db 2036 ATCATTTAGAGTCTCC 2050

RESULT 27
ACC60521
ID ACC60521 standard; cDNA; 5111 BP.
XX ACC60521;
AC 19-JUN-2003 (first entry)
XX
XX Polynucleotide relating to the invention SEQ ID NO: 41.
DE Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic;
XX antiproliferative; cardiant; cytosolic; gene therapy; liver disease;
XX proliferative disorder; renal failure; cardiovascular disorder;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
OS Homo sapiens.
XX
XX MO200257460-A2.
XX
XX 25-JUL-2002.
XX
XX 20-DEC-2001; 2001MO-US050459.
XX
XX 20-DEC-2001; 2000US-0256868P.
XX 30-MAR-2001; 2001US-0280186P.
XX 01-MAY-2001; 2001US-0287735P.
XX 05-JUN-2001; 2001US-0295848P.
XX 25-JUN-2001; 2001US-0300465P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
XX Stewers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
XX Krystek S, Mcatee P, Suchard S, Banas D;
XX
XX MPI; 2002-599721/64.
XX P-PSDB; ABR52352.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
XX the prevention or treatment of e.g. proliferative and cardiovascular
XX disorders.
XX
XX Claim 1; Fig 5; 801pp; English.

CC The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antarthritic, antiproliferative, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention

XX Sequence 5111 BP; 1413 A; 1112 C; 1232 G; 1354 T; 0 U; 0 Other;

Query Match 89.8%; Score 1791; DB 6; Length 5111;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATGAACTCAATTTGTTACTGAGAGGTGGCTGTCTGGAA 60
Db 470 ATGGCCCATGAGATGATGAACTCAATTTGTTACTGAGAGGTGGCTGTCTGGAA 529
QY 61 AGTGAACGGAAAAAGTGTCTGCTAATTGATAGCCGGCCATTGTTGAAATACATATCC 120
Db 530 AGTGAACGGAAAAAGTGTCTGCTAATTGATAGCCGGCCATTGTTGAAATACATATCC 589
QY 121 CACATTTTGGAAAGCCATTAATATCAATGCTCCAGCTTATGAGGAAAGTTGCAACG 180
Db 590 CACATTTTGGAAAGCCATTAATATCAATGCTCCAGCTTATGAGGAAAGTTGCAACG 649
QY 181 GACAAAGTTTAATTACAGAGCTCATCCAGATTCCAGCAACATTAAGTTGACATTGAT 240
Db 650 GACAAAGTTTAATTACAGAGCTCATCCAGATTCCAGCAACATTAAGTTGACATTGAT 709
QY 241 TGCAGTCAGAGAGTTGATGTTACGATCAAGTCCCAAGATGTTGCTCTCTCTTCA 300
Db 710 TGCAGTCAGAGAGTTGATGTTACGATCAAGTCCCAAGATGTTGCTCTCTCTTCA 769
QY 301 GACTGTTTCTGACTGTACTTCTGGGTAACTGAGAGAGTTCAACTGTGTCACCTG 360
Db 770 GACTGTTTCTGACTGTACTTCTGGGTAACTGAGAGAGTTCAACTGTGTCACCTG 829
QY 361 CTTCAGAGTGGGTTTGTGCTGAGTCTCTGTTTTCCTGGGCTCTGTGAGAGAAATCC 420
Db 830 CTTCAGAGTGGGTTTGTGCTGAGTCTCTGTTTTCCTGGGCTCTGTGAGAGAAATCC 889
QY 421 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGCTTACTGTTGCCAATTTGGCCCAACC 480
Db 890 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTGCTTACTGTTGCCAATTTGGCCCAACC 949
QY 481 CGAATTCCTCCCAATTTTATCTGGCTGCGAGCGAGAGTCTCTCAACAGAGCTGATA 540
Db 950 CGAATTCCTCCCAATTTTATCTGGCTGCGAGCGAGAGTCTCTCAACAGAGCTGATA 1009
QY 541 CAGCAGAAATGGATTTGTTATGTTAAATGACAGCTATACCTGTCCAAAGCCTGACTT 600
Db 1010 CAGCAGAAATGGATTTGTTATGTTAAATGACAGCTATACCTGTCCAAAGCCTGACTT 1069
QY 601 ATCCCGAGTCTCATTTCTGCGGTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 660
Db 1070 ATCCCGAGTCTCATTTCTGCGGTGCTGCTGATGACAGCTTTTGTGAGAAATTTTG 1129
QY 661 CCGTGTGGAGAAATACATTAATTTCAATTGAGAAAGCAAAAGCCCTCCAAATGATGTT 720
Db 1130 CCGTGTGGAGAAATACATTAATTTCAATTGAGAAAGCAAAAGCCCTCCAAATGATGTT 1189
QY 721 CTATGTCATGTTTATGCTGGAGATCTCCGCTCCGCCACCATGCTATACATCATG 780
Db 1190 CTATGTCATGTTTATGCTGGAGATCTCCGCTCCGCCACCATGCTATACATCATG 1249
QY 781 AAGAGATGAGACATGCTTTAGATGAAAGCTTACAGATTTGTGAGAAAGAAAGACTTACT 840
Db 1250 AAGAGATGAGACATGCTTTAGATGAAAGCTTACAGATTTGTGAGAAAGAAAGACTTACT 1309
QY 841 AATATCTCCAAATTTCAATTTTCTGGGCCAACTCTGGACTATGAGAAAGATTAAGAAC 900
Db 1310 AATATCTCCAAATTTCAATTTTCTGGGCCAACTCTGGACTATGAGAAAGATTAAGAAC 1369
QY 901 CAGACTGAGCATCAGGAGGCAAGAGCAAACTCAAGCTGCTCAGCTGAGAGAGCCAAAT 960
Db 1370 CAGACTGAGCATCAGGAGGCAAGAGCAAACTCAAGCTGCTCAGCTGAGAGAGCCAAAT 1429
QY 961 GAACCTGTCTCTGCTGTCTCAAGAGGTGAGCAAGAAAGCGAGAGCCCTCAGTCCACCC 1020
Db 1430 GAACCTGTCTCTGCTGTCTCAAGAGGTGAGCAAGAAAGCGAGAGCCCTCAGTCCACCC 1489
QY 1021 TGTGCGAATCTGTCTAATCTCAAGAGGACAGAGCAAAAGCCCGTGTGATCCGCCACGCTG 1080


```

Db      1490  TGTGCGCACTGTACTCTCAGAGGACAGAGCAAAAGCCCGTGCAATCCCGCCAGGCTG 1549
QY      1081  CCCAGCGGCCAGGTCAGACCCGTCGTGTTAAGGACAGCCCGCTGTTAAGAGCGCTC 1140
Db      1550  CCCAGCGGCCAGGTCAGACCCGTCGTGTTAAGGACAGCCCGCTGTTAAGAGCGCTC 1609
QY      1141  AGTGGGCTGCACTGTCCGAGACAGGCTGGAGAGACAGCAATTAAGCTCAAGCGTCTC 1200
Db      1610  AGTGGGCTGCACTGTCCGAGACAGGCTGGAGAGACAGCAATTAAGCTCAAGCGTCTC 1669
QY      1201  TCTCTGGAATCAATCAATGTTTCAATTAATGACCAATGAGAGATCCCTTAATGAGCTT 1260
Db      1670  TCTCTGGAATCAATCAATGTTTCAATTAATGACCAATGAGAGATCCCTTAATGAGCTT 1729
QY      1261  TCCTCATAGAAAGTCTTGGAAATACAAACCTTCCACTACTCTGATGGAGCAAC 1320
Db      1730  TCCTCATAGAAAGTCTTGGAAATACAAACCTTCCACTACTCTGATGGAGCAAC 1789
QY      1321  AAGCTATGCCAGTCTCCCTGTTCAAGAACTATCGAGAGACTCCCGAAACAGTCTCT 1380
Db      1790  AAGCTATGCCAGTCTCCCTGTTCAAGAACTATCGAGAGACTCCCGAAACAGTCTCT 1849
QY      1381  GATTAAGAGAGAAAGCCACATCCCAAGAACTGACAGCCGCAAGCCTTCAAGACCAAG 1440
Db      1850  GATTAAGAGAGAAAGCCACATCCCAAGAACTGACAGCCGCAAGCCTTCAAGACCAAG 1909
QY      1441  AGCAAGGATTTGATTTGGTCAAGAACAGAGAGAGTGGACCGCCAGAGGTCTCTTTA 1500
Db      1910  AGCAAGGATTTGATTTGGTCAAGAACAGAGAGAGTGGACCGCCAGAGGTCTCTTTA 1969
QY      1501  TCTCCACTGCATGAAGTGGAGCGTGGAGAGCAATTAACCAACAGCTCTCTTTTCCGAC 1560
Db      1970  TCTCCACTGCATGAAGTGGAGCGTGGAGAGCAATTAACCAACAGCTCTCTTTTCCGAC 2029
QY      1561  CTTTCCACGAGCAGACGACCTCAAGAACTGCTGCGCTGGGCTTTAAGGCTTGGCAC 1620
Db      2030  CTTTCCACGAGCAGACGACCTCAAGAACTGCTGCGCTGGGCTTTAAGGCTTGGCAC 2089
QY      1621  TCGGATATCTTTGGCCCCCAGACCTTACCCCTGACGACGAGCTGTATTTTGGC 1680
Db      2090  TCGGATATCTTTGGCCCCCAGACCTTACCCCTGACGACGAGCTGTATTTTGGC 2149
QY      1681  ACAAGATCCTCACTTAATCTGCTCAAGCATATCAAGAGGACAGTGCATTAATCT 1740
Db      2150  ACAAGATCCTCACTTAATCTGCTCAAGCATATCAAGAGGACAGTGCATTAATCT 2209
QY      1741  GCCTACAGCTGACGACGCTGCCACTTTCGAGAGACCAAGTCTATTTCTGTGGCAGGGG 1800
Db      2210  GCCTACAGCTGACGACGCTGCCACTTTCGAGAGACCAAGTCTATTTCTGTGGCAGGGG 2269
QY      1801  CAGAAAGCCAGTGAACAGAGCTTGAATTCGCGCGAGAGCTGGCATGAAGAGCCCTTTGA 1860
Db      2270  CAGAAAGCCAGTGAACAGAGCTTGAATTCGCGCGAGAGCTGGCATGAAGAGCCCTTTGA 2329
QY      1861  AAGCATTTAAACGAGAAAGTCCCAATTAATTAAGAGAGATCAATGTCAGAGAAC 1920
Db      2330  AAGCATTTAAACGAGAAAGTCCCAATTAATTAAGAGAGATCAATGTCAGAGAAC 2389
QY      1921  AGGTACGAGAAAGAGCTGGGAAAGTGGAGAGTCAAGTCTTAAGCTTTGGGAGAGATGAA 1980
Db      2390  AGGTACGAGAAAGAGCTGGGAAAGTGGAGAGTCAAGTCTTAAGCTTTGGGAGAGATGAA 2449
QY      1981  ATCATTTAGGCTTCC 1995
Db      2450  ATCATTTAGGCTTCC 2464

```

```

RESULT 28
AAF30479
ID      AAF30479 standard; cDNA, 2118 BP.
XX
AC      AAF30479;

```

```

XX      29-MAY-2001 (first entry)
DT
XX      Human protein phosphatase and kinase protein-4 cDNA 1234795CB1.
DE
XX
XX      Protein phosphatase and kinase protein; PPHKP-4; human;
XX      gastrointestinal disorder; immune system disorder; neurological disorder;
XX      cell proliferative disorder; cancer; diagnosis; therapy; ss.
OS
XX      Homo sapiens.
XX
XX      Key      Location/Qualifiers
XX      CDS      25..2025
XX              /*tag= a
XX              /*tag= b
XX              /*tag= c
XX              /note= "unique fragment"
XX              /note= "unique fragment"
XX              /note= "unique fragment"
XX
XX      WO200120004-A2.
XX
XX      22-MAR-2001.
XX
XX      14-SEP-2000; 2000WO-US025515.
XX
XX      15-SEP-1999; 99US-0154141P.
XX
XX      (INCY-) INCYTE GENOMICS INC.
XX
XX      Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y,
XX      Lu DAM;
XX      WPI: 2001-244811/25.
XX      P-PSDB; AAB20325.
XX
XX      Novel human protein phosphatase and kinase proteins for diagnosis,
XX      treatment and prevention of gastrointestinal, immune system, neurological
XX      and cell proliferative disorders.
XX
XX      Claim 5; Page 97-98; 103pp; English.
XX
XX      The present sequence is that of cDNA encoding novel human protein
XX      phosphatase and kinase protein PPHKP-4 (see AAB20325). The cDNA was
XX      initially identified in Incyte Clone ID No. 1234795CB1, from a foetal
XX      lung tissue library. Tissues that express PPHKP-4 (as a fraction of total
XX      tissues expressing PPHKP-4) include gastrointestinal (0.385),
XX      cardiovascular (0.154), nervous (0.154) and reproductive (0.154).
XX      Diseases or conditions associated with tissues expressing PPHKP-4 (as a
XX      fraction of total tissues expressing PPHKP-4) include cancer (0.692),
XX      inflammation or trauma (0.308) and cell proliferation (0.231). The
XX      encoded protein shows homology to mouse neuronal tyrosine threonine
XX      phosphatase 1. The invention provides human PPHKP-1 to -11 polypeptides
XX      (see AAB20322-32) and polynucleotides (see AAF30476-86). It also provides
XX      expression vectors, host cells, antibodies, agonists and antagonists, as
XX      well as methods for diagnosing, treating or preventing disorders
XX      associated with expression of PPHKP, including gastrointestinal
XX      disorders, immune system disorders, neurological disorders and cell
XX      proliferative disorders, including cancer
XX
XX      Sequence 2118 BP; 550 A; 552 C; 506 G; 510 T; 0 U; 0 Other;
SQ

```

```

Query Match      79.8%; Score 1593; DB 4; Length 2118;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

```

```

QY      1  ATGGCCCATGAGATGATGGAATCAATTTGTTACTGAGAGGTTGTGGCTGTGCTGGA 60
Db      25  ATGGCCCATGAGATGATGGAATCAATTTGTTACTGAGAGGTTGTGGCTGTGCTGGA 84
QY      61  AGTGAACGGAAGAAAGTGTGCTTAATGATAGCCGCAATTTGTGAAATCAATCATCC 120

```


Db 85 AGTGAACGGAAGAGTGTCTAATGATAGCCGGCATTGTGGAATACATATACCTC 144
QY 121 CAAATTTTGGAAACCAATTAATCACTGCTCCAGTTTATGAAGGAAGTTGCAACG 180
Db 145 CAAATTTTGGAAACCAATTAATCACTGCTCCAGTTTATGAAGGAAGTTGCAACG 204
QY 181 GACAAAGTGTAAATTCAGAGCTCATCCAGCAATTCAGCAAAATTAAGTTGACATTTG 240
Db 205 GACAAAGTGTAAATTCAGAGCTCATCCAGCAATTCAGCAAAATTAAGTTGACATTTG 264
QY 241 TGCAGTCAGAAAGTTGATTTAAGATCAAAAGTCCCAAGATTTGCTCTCTCTCA 300
Db 265 TGCAGTCAGAAAGTTGATTTAAGATCAAAAGTCCCAAGATTTGCTCTCTCTCA 324
QY 301 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAAGAGCTTCAACTCTGTTACCTG 360
Db 325 GACTGTTTCTCACTGATCTTCTGGGTAACTGGAAGAGCTTCAACTCTGTTACCTG 384
QY 361 CTTCAGAGTGGGTTTGTGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAGGAATCC 420
Db 385 CTTCAGAGTGGGTTTGTGAGTCTCTGTTGTTTCCCTGGCTCTGTGAAGGAATCC 444
QY 421 ACTCTAGTCCCTAATCTGATTTTCTAGCCTTGTCTTACTGTTGCAATTTGGCCAC 480
Db 445 ACTCTAGTCCCTAATCTGATTTTCTAGCCTTGTCTTACTGTTGCAATTTGGCCAC 504
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCGACAGAGATGCTCTCAACAGAGCTGAT 540
Db 505 CGAATTTCTCCCAATCTTTATCTTGGCTGCGACAGAGATGCTCTCAACAGAGCTGAT 564
QY 541 CACAGAAATGGATGGTGTATGTATTAATGCCAGCTATCTGTCCAAAGCTGACTT 600
Db 565 CACAGAAATGGATGGTGTATGTATTAATGCCAGCTATCTGTCCAAAGCTGACTT 624
QY 601 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGGAATTTTG 660
Db 625 ATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGGAATTTTG 684
QY 661 CCGTGTGTGGAACAAATCAGTGAATTTCAATTGAAAGAAAGCAAGCTTCAATGATG 720
Db 685 CCGTGTGTGGAACAAATCAGTGAATTTCAATTGAAAGAAAGCAAGCTTCAATGATG 744
QY 721 CTAGTGACATGTTAGTGGGATCTCCGCTCCGCCACATGCTATGCTATCATCATG 780
Db 745 CTAGTGACATGTTAGTGGGATCTCCGCTCCGCCACATGCTATGCTATCATCATG 804
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 840
Db 805 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 864
QY 841 ATATCTCCAAATCTCAATTTTCTGGGCAATCTCTGATCTATGAAGAAAGTTAAGAC 900
Db 865 ATATCTCCAAATCTCAATTTTCTGGGCAATCTCTGATCTATGAAGAAAGTTAAGAC 924
QY 901 CAACTGAGAGCATCAGGGCCAAAGAGCAATCAAGTGTGACCTGAGAGGAAGCAAT 960
Db 925 CAACTGAGAGCATCAGGGCCAAAGAGCAATCAAGTGTGACCTGAGAGGAAGCAAT 984
QY 961 GAACTGTCTCTGCTCTCAGAGGTTGACAGAAAAAGCAGAGCCCTCAGTCCACCC 1020
Db 985 GAACTGTCTCTGCTCTCAGAGGTTGACAGAAAAAGCAGAGCCCTCAGTCCACCC 1044
QY 1021 TGTGCGAGCTCTGCTCTCAGAGGTTGACAGAAAAAGCAGAGCCCTCAGGCGG 1080
Db 1045 TGTGCGAGCTCTGCTCTCAGAGGTTGACAGAAAAAGCAGAGCCCTCAGGCGG 1104
QY 1081 CCCAGGCTGCGAGCGTGCAGCGCTGTGTTAGAGGACAGCCGCTGTTAGAGGCGCTC 1140
Db 1105 CCCAGGCTGCGAGCGTGCAGCGCTGTGTTAGAGGACAGCCGCTGTTAGAGGCGCTC 1164
QY 1141 AGTGGGCTGACCTGTCTCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1200
Db 1165 AGTGGGCTGACCTGTCTCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCTTCTTC 1224

QY 1201 TCTCTGATATCAAAATCAGTTTCATATTCAGCCAGACAGTGGACATCTTACATGGCTTC 1260
Db 1225 TCTCTGATATCAAAATCAGTTTCATATTCAGCCAGATGGACATCTTACATGGCTTC 1284
QY 1261 TCTCTATCAGAAAGTCTTTTGAATTAATCAAACTTTCATCTGTGATGGACCAAC 1320
Db 1285 TCTCTATCAGAAAGTCTTTTGAATTAATCAAACTTTCATCTGTGATGGACCAAC 1344
QY 1321 AAGCTATGCGAAGTTCTCCCTGTTACAGAACTATCGAGACAGCTCCGAAACAGT--C 1378
Db 1345 AAGCTATGCGAAGTTCTCCCTGTTACAGAACTATCGAGACAGCTCCGAAACAGT--C 1404
QY 1379 CTGATTAAGGAGGAGCAGATCCCAAGAAAGTGCAGACCGGACAGGCTTTCAGACAGC 1438
Db 1405 CTGATTAAGGAGGAGCAGATCCCAAGAAAGTGCAGACCGGACAGGCTTTCAGACAGC 1464
QY 1439 AGAGCAAGGATTCATTCGCTCAGAACAGAGCAGTGGCAACCGCCAGAGGTCCTTT 1498
Db 1465 AGAGCAAGGATTCATTCGCTCAGAACAGAGCAGTGGCAACCGCCAGAGGTCCTTT 1524
QY 1499 TATCTCACTGATCGAAGTGGAGCGTGGAGACAAATTAACACACAGCTTCTTTTG 1558
Db 1525 TATCTCACTGATCGAAGTGGAGCGTGGAGACAAATTAACACACAGCTTCTTTTG 1584
QY 1559 GCCTTCCACAGCCAGCAGACCTTCAGAAAGTCTGCTGGGCTGGG--CCTTAAAGGCTGG 1617
Db 1585 GCCTTCCACAGCCAGCAGACCTTCAGAAAGTCTGCTGGGCTGGG--CCTTAAAGGCTGG 1644
QY 1618 CACTCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTATTTT 1677
Db 1645 CACTCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTATTTT 1704
QY 1678 GCCACAGATCTTACACTTTCTACTGCTCTAGCCATCTACGAGAGCAGTGCAGTAC 1737
Db 1705 GCCACAGATCTTACACTTTCTACTGCTCTAGCCATCTACGAGAGCAGTGCAGTAC 1764
QY 1738 TCTGCTTACAGTGCAGCAGCAGCTGCCACTTGGGAGCAAGTCTATCTGTGGGAGG 1797
Db 1765 TCTGCTTACAGTGCAGCAGCAGCTGCCACTTGGGAGCAAGTCTATCTGTGGGAGG 1824
QY 1798 CCGCAGAAACCAAGTACAGAGCTGACTCGCGCGAGCTGCGATGAAGAGAGCCCTTT 1857
Db 1825 CCGCAGAAACCAAGTACAGAGCTGACTCGCGCGAGCTGCGATGAAGAGAGCCCTTT 1884
QY 1858 GAAAGCAGTTTAAACGCAAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTACAG 1917
Db 1885 GAAAGCAGTTTAAACGCAAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTACAG 1944
QY 1918 AACAGTCAACGGGAAGAGCTGGGGAAGTGGGCAAGTCAATGCTTTTCCGGCAGCATG 1977
Db 1945 AACAGTCAACGGGAAGAGCTGGGGAAGTGGGCAAGTCAATGCTTTTCCGGCAGCATG 2004
QY 1978 GAAATCATTAAGGCTCC 1995
Db 2005 GAAATCATTAAGGCTCC 2022

Search completed: June 22, 2004, 10:38:26
Job time : 1150 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2004, 10:19:15 ; Search time 1174 Seconds

(without alignments)
7784.730 Million cell updates/sec

Title: US-10-029-345A-108_COPY_538_2532

Perfect score: 1995
Sequence: 1 atggcccatgagatgattgtg.....tggaaatcatctgagctcc 1995

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3017426 seqs, 2290544650 residues

Word size : 1419

Total number of hits satisfying chosen parameters: 12

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database :

Published Applications NA:
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1893	94.9	1998	9 US-09-816-494-3	Sequence 3, Appl
2	1893	94.9	1998	16 US-10-377-072-27	Sequence 27, Appl
3	1893	94.9	3059	17 US-10-257-026-1	Sequence 1, Appl
4	1893	94.9	3496	9 US-09-964-277-1	Sequence 1, Appl
5	1893	94.9	3544	9 US-09-816-494-1	Sequence 1, Appl
6	1893	94.9	3544	16 US-10-377-072-25	Sequence 25, Appl
7	1893	94.9	3625	13 US-10-425-114-26234	Sequence 17, Appl
8	1893	94.9	3766	13 US-10-343-357-17	Sequence 115, Appl
9	1893	94.9	4790	17 US-10-648-593-115	Sequence 520, Appl
10	1849	92.7	2966	13 US-10-296-115-520	Sequence 2, Appl
11	1842	92.3	2732	13 US-10-168-506-2	Sequence 673, Appl
12	1791	89.8	2102	16 US-10-094-749-673	Sequence 673, Appl

ALIGNMENTS

RESULT 1
US-09-816-494-3
; Sequence 3, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-03002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIORITY FILING DATE: 2001-03-23
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-816-494-3

Query Match 94.9%; Score 1893; DB 9; Length 1998;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGGCCCATGAGATGATGGAACACTCAATTTGTTAGAGAGTTGGCTGTGCGAA	60
DB	1	ATGGCCCATGAGATGATGGAACACTCAATTTGTTAGAGAGTTGGCTGTGCGAA	60
QY	61	AGTGAAGCGGAAAAATGCTGCTTAATTGATAGCCGCCATTTGTGGAATACATATCC	120
DB	61	AGTGAAGCGGAAAAATGCTGCTTAATTGATAGCCGCCATTTGTGGAATACATATCC	120
QY	121	CACATTTTGAAGCCATTAATATCACTGCTCCAACTTATGAGGAAAGTTGCAACAG	180
DB	121	CACATTTTGAAGCCATTAATATCACTGCTCCAACTTATGAGGAAAGTTGCAACAG	180
QY	181	GACAAAGTTTATTAACAAGCTCATCAAGATTCAGCGAAACATTAAGTTGACATGAT	240
DB	181	GACAAAGTTTATTAACAAGCTCATCAAGATTCAGCGAAACATTAAGTTGACATGAT	240
QY	241	TGACGTCAGAAAGTTGATTAATCAATCAAGTCCCAAGATTTGCTCTCTCTCA	300
DB	241	TGACGTCAGAAAGTTGATTAATCAATCAAGTCCCAAGATTTGCTCTCTCTCA	300
QY	301	GACTGTTTCTACCTGACTCTGGGTAAACTGAGAAAGCTTCAACTCTGTCACTG	360
DB	301	GACTGTTTCTACCTGACTCTGGGTAAACTGAGAAAGCTTCAACTCTGTCACTG	360
QY	361	CTTGACAGTGGTTTGTGCTGATCTCTGTTTCTGCTGCTCTGTGAGAAATCC	420
DB	361	CTTGACAGTGGTTTGTGCTGATCTCTGTTTCTGCTGCTCTGTGAGAAATCC	420
QY	421	ACTTATGTCCTACCTGCAATTTCTGAGCTTCTCTTACCTGTTCCAACTGGCCAAAC	480
DB	421	ACTTATGTCCTACCTGCAATTTCTGAGCTTCTCTTACCTGTTCCAACTGGCCAAAC	480
QY	481	CGAATCTTCCCAATTTTATCTTGCTGCGACGAGATGCTCTCAACAGAGCTGATG	540
DB	481	CGAATCTTCCCAATTTTATCTTGCTGCGACGAGATGCTCTCAACAGAGCTGATG	540
QY	541	CAGCAAAATGAGATTTGATGTTAAATGACGATATACCTGTCCAAAGCTGACTTT	600
DB	541	CAGCAAAATGAGATTTGATGTTAAATGACGATATACCTGTCCAAAGCTGACTTT	600
QY	601	ATCCCGAAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	660
DB	601	ATCCCGAAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	660
QY	661	CGGTGTTGGAACAATCAGATTTCAATTGAGAAAGCAAAAGCTTCAATGATGTT	720

Db 661 CCGTGGTGGACAAATCAGTATGATTTTCATTTGAGAAAAGAAAAGCCCTCCAAATGAGATGTT 720
Qy 721 CTAAGTCACTGTTTAGCTGGGATCTCCCGCTCCGACCATCGTATGCGCTCATCATG 780
Db 721 CTAAGTCACTGTTTAGCTGGGATCTCCCGCTCCGACCATCGTATGCGCTCATCATG 780
Qy 781 AAGAGATGAGATGCTTTAGATGAAGCTTACATTTTGGAAAAGAAAAGACTTACT 840
Db 781 AAGAGATGAGATGCTTTAGATGAAGCTTACATTTTGGAAAAGAAAAGACTTACT 840
Qy 841 AATATCCAACTTCATTTTCTGGGCCAACTCTGATGAGAAAGATTAAAGAC 900
Db 841 AATATCCAACTTCATTTTCTGGGCCAACTCTGATGAGAAAGATTAAAGAC 900
Qy 901 CAGACTGGAGATCAAGGGCCAAAGGCAAACTCAAGCTGTGCACTGAGAAAGCCAAAT 960
Db 901 CAGACTGGAGATCAAGGGCCAAAGGCAAACTCAAGCTGTGCACTGAGAAAGCCAAAT 960
Qy 961 GAACCTGTCCCTGTCTCTCAGAGGGGTGAGACAAAAGCGAGCCCTCAGTCCACCC 1020
Db 961 GAACCTGTCCCTGTCTCTCAGAGGGGTGAGACAAAAGCGAGCCCTCAGTCCACCC 1020
Qy 1021 TGTGCGCACTCTGCTACTCTCAGAGGCGAGAGCAAAAGCCCTGATCCCGCAGCGTG 1080
Db 1021 TGTGCGCACTCTGCTACTCTCAGAGGCGAGAGCAAAAGCCCTGATCCCGCAGCGTG 1080
Qy 1081 CCCAGCGGCGCCAGGCGTGCAGGCGGTCTGTTAAGAGACCCCGCTGTTACAGGGCTC 1140
Db 1081 CCCAGCGGCGCCAGGCGTGCAGGCGGTCTGTTAAGAGACCCCGCTGTTACAGGGCTC 1140
Qy 1141 AGTGGGCTGCAGCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCCTTCTTC 1200
Db 1141 AGTGGGCTGCAGCTGTCCGAGACAGGCTGGAAGACAGCAATTAAGCTCAAGCCTTCTTC 1200
Qy 1201 TCTGTGATATCAAAATGATTTTCAATTTAGCCAGCATGCGACATCTTACATGCTTC 1260
Db 1201 TCTGTGATATCAAAATGATTTTCAATTTAGCCAGCATGCGACATCTTACATGCTTC 1260
Qy 1261 TCCTCATCAGAAAGTCTTTGGAATACTAACAACCTTCCACTCTGAGTGGAGCAAC 1320
Db 1261 TCCTCATCAGAAAGTCTTTGGAATACTAACAACCTTCCACTCTGAGTGGAGCAAC 1320
Qy 1321 AAGCTATGCGAGTTTCCCTGTTCAAGAACTATCCGAGACAGCTCCGAAACAGTCTCT 1380
Db 1321 AAGCTATGCGAGTTTCCCTGTTCAAGAACTATCCGAGACAGCTCCGAAACAGTCTCT 1380
Qy 1381 GATTAAGAGAGAACCCAGCATCCCAAGAAAGCTGCAGACCGCGCTTCAACAGCCAG 1440
Db 1381 GATTAAGAGAGAACCCAGCATCCCAAGAAAGCTGCAGACCGCGCTTCAACAGCCAG 1440
Qy 1441 AGCAAGCGATTGATTCGGTCAAGACCAAGCAAGTGCAGACCGCCAGAGTCCCTTTTA 1500
Db 1441 AGCAAGCGATTGATTCGGTCAAGACCAAGCAAGTGCAGACCGCCAGAGTCCCTTTTA 1500
Qy 1501 TCTCCACTGATGAAGTGGAGCGCTGGAGCAATTAACAACAAGCTCTCTTTGGGC 1560
Db 1501 TCTCCACTGATGAAGTGGAGCGCTGGAGCAATTAACAACAAGCTCTCTTTGGGC 1560
Qy 1561 CTTTCACAGCAGCAGACCACTCAAGAAAGTGTGCTGCGCTTGAAGGCTGGGAC 1620
Db 1561 CTTTCACAGCAGCAGACCACTCAAGAAAGTGTGCTGCGCTTGAAGGCTGGGAC 1620
Qy 1621 TCGGATATCTGGCCCCCGACCTCTACCCCTTCCCTGACCAAGAGCTGTATTTGGC 1680
Db 1621 TCGGATATCTGGCCCCCGACCTCTACCCCTTCCCTGACCAAGAGCTGTATTTGGC 1680
Qy 1681 ACAGAGTCTCACTTCACTGCTCAGCAATCTTAAGAGAGAGCCAGTTACTCT 1740
Db 1681 ACAGAGTCTCACTTCACTGCTCAGCAATCTTAAGAGAGAGCCAGTTACTCT 1740
Qy 1741 GCCTACAGCTGCAGCCAGCTGCCCACTTGCGGAGAACAAAGTCTATTTCTGTCGAGCGG 1800
Db 1741 GCCTACAGCTGCAGCCAGCTGCCCACTTGCGGAGAACAAAGTCTATTTCTGTCGAGCGG 1800

Db 1741 GCCTACAGCTGCAGCCAGCTGCCCACTTGCGGAGAACAAAGTCTATTTCTGTCGAGCGG 1800
Qy 1801 CAGAAGCGAAGTACAGAGCTGACTGCGCGCGAGAGTGGCATGAAGAGCCCTTTGAA 1860
Db 1801 CAGAAGCGAAGTACAGAGCTGACTGCGCGCGAGAGTGGCATGAAGAGCCCTTTGAA 1860
Qy 1861 AAGCAGTTTAAAGCAGAGCTGCCAAATGAAATTTGGAGAGAGCATGTCAGAGAAC 1920
Db 1861 AAGCAGTTTAAAGCAGAGCTGCCAAATGAAATTTGGAGAGAGCATGTCAGAGAAC 1920
Qy 1921 AGGTCCGGGAAGAGTGGGGAAGTGGGAGTCACTTAAGCTTTTGGGCGAGATGANA 1980
Db 1921 AGGTCCGGGAAGAGTGGGGAAGTGGGAGTCACTTAAGCTTTTGGGCGAGATGANA 1980
Qy 1981 ATCATTGAGGTCTCC 1995
Db 1981 ATCATTGAGGTCTCC 1995

541-1945P 1455

RESULT 2
US-10-377-072-27
Sequence 27, Application US/10377072
Publication No. US20040009501A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals Inc.
APPLICANT: Curtis, Rory A.J.
APPLICANT: Logan, Thomas Joseph
APPLICANT: Glucksmann, Maria A.
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Chun, Miyoung
APPLICANT: Tsai, Fong-Ying
TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
FILE REFERENCE: MP103-0180NMIM
CURRENT APPLICATION NUMBER: US/10/377, 072
PRIOR APPLICATION NUMBER: US 09/895,860
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215,370
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/723,806
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 60/187,455
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/843,297
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199,801
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: US 09/861,801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205,508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/816,494
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/815,419
PRIOR FILING DATE: 2001-03-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 1998
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1998)
US-10-377-072-27

Query Match 94.9%; Score 1893; DB 16; Length 1998;
Best Local Similarity 99.9%; Pred. No. 0;

FILE REFERENCE: DUSP10K.DMS
CURRENT APPLICATION NUMBER: US/10/257,026
CURRENT FILING DATE: 2003-11-07
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 3059
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (127) .. (2121)
US-10-257-026-1

Query Match 94.9%; Score 1893; DB 17; Length 3059;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGCCCATGAGATGATGGAACTCAAAATGTTACTGAGAGTTGGTGGCTCTGCGAA 60
DB 127 ATGGCCCATGAGATGATGGAACTCAAAATGTTACTGAGAGTTGGTGGCTCTGCGAA 186
QY 61 AGTGAACGGAAGAAAGTGTCTGTAATGATAGCCGCGCATTTGTGAATACATATCC 120
DB 187 AGTGAACGGAAGAAAGTGTCTGTAATGATAGCCGCGCATTTGTGAATACATATCC 246
QY 121 CACATTTTGAAGCCATTAATATCACTGCTCCAAAGCTTATGAAGCGAAGTTGCAACG 180
DB 247 CACATTTTGAAGCCATTAATATCACTGCTCCAAAGCTTATGAAGCGAAGTTGCAACG 306
QY 181 GACAAAGTGTAAATTAAGAGCTATCCAGCATTCAGCGAAACATTAAGTTGACATTTGAT 240
DB 307 GACAAAGTGTAAATTAAGAGCTATCCAGCATTCAGCGAAACATTAAGTTGACATTTGAT 366
QY 241 TGCATGAGAAAGTTGATGTTTACGATCAAAAGTCCCAAGATGTGGCTCTCTCTTCA 300
DB 367 TGCATGAGAAAGTTGATGTTTACGATCAAAAGTCCCAAGATGTGGCTCTCTCTTCA 426
QY 301 GACTGTTTCTCACTGTACTTCTGGGTAACTGGAAGAGCTTCAACTCTGTTCACTG 360
DB 427 GACTGTTTCTCACTGTACTTCTGGGTAACTGGAAGAGCTTCAACTCTGTTCACTG 486
QY 361 CTGTCAGAGTGGGTTTGCTGAGTCTCTCCGTTGTTTCCCTGGCTCTGTGAAGAAATCC 420
DB 487 CTGTCAGAGTGGGTTTGCTGAGTCTCTCCGTTGTTTCCCTGGCTCTGTGAAGAAATCC 546
QY 421 ACTCTAGTCCCTAAGTCACTTCTGAGCTTCTGCTTCACTGTTGCCAACTTGGGCCAAC 480
DB 547 ACTCTAGTCCCTAAGTCACTTCTGAGCTTCTGCTTCACTGTTGCCAACTTGGGCCAAC 606
QY 481 CGAATCTTCCCAATCTTTATCTGGCTGCGCAGCGAATGTCCTCAACAGAGAGCTGATA 540
DB 607 CGAATCTTCCCAATCTTTATCTGGCTGCGCAGCGAATGTCCTCAACAGAGAGCTGATA 666
QY 541 CAGCAGAAATGGAGTGTATATGTTAAATGCAAGTATACCTGTCCAAAGCTTACCTTT 600
DB 667 CAGCAGAAATGGAGTGTATATGTTAAATGCAAGTATACCTGTCCAAAGCTTACCTTT 726
QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 660
DB 727 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 786
QY 661 CCGTGGTTTGGACAATACGTATATTTCAATGGAAGAAAGCAAGCTCCATAGATGTGT 720
DB 787 CCGTGGTTTGGACAATACGTATATTTCAATGGAAGAAAGCAAGCTCCATAGATGTGT 846
QY 721 CTAGTGAATGTTTGTAGTGGATCTCCGCTCCGCAACATCGCTATAGGCTTATATATG 780
DB 847 CTAGTGAATGTTTGTAGTGGATCTCCGCTCCGCAACATCGCTATAGGCTTATATATG 906
QY 781 AAGAGATGAGCATGTCTTTTGAATGAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 840
DB 907 AAGAGATGAGCATGTCTTTTGAATGAGCTTACAGATTTGTGAAGAAAGAAAGCACTACT 966
```

```
QY 841 ATATCTCAAACTTCAATTTTCTGGGCCAACTCTGGAATATAGAAAGATTAAAGAC 900
DB 967 ATATCTCAAACTTCAATTTTCTGGGCCAACTCTGGAATATAGAAAGATTAAAGAC 1026
QY 901 CAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTGTCTGACCTGAGAGCAAAAT 960
DB 1027 CAGACTGAGCATCAGGGCCAAAGCAAACTCAAGCTGTCTGACCTGAGAGCAAAAT 1086
QY 961 GAACTGTCCCTGTCTCTCAGAGGTGGACGAAAGCAAGAGCCCTCTCACTCACCC 1020
DB 1087 GAACTGTCCCTGTCTCTCAGAGGTGGACGAAAGCAAGAGCCCTCTCACTCACCC 1146
QY 1021 TGTGCGACTGTCTTACTCAGAGGCAAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1080
DB 1147 TGTGCGACTGTCTTACTCAGAGGCAAGAGCAAAAGCCCGTGCATCCCGCAGCGTG 1206
QY 1081 CCCAGCGTCCCGAGGTGACGCGCTGCTGTTAAGAGCAGCCCGTGTACAGGCGCTC 1140
DB 1207 CCCAGCGTCCCGAGGTGACGCGCTGCTGTTAAGAGCAGCCCGTGTACAGGCGCTC 1266
QY 1141 AGTGGCTGCACTGTCCGACAGACAGGCTGGAAGCAATAGCTCAAGCTTCTCTTC 1200
DB 1267 AGTGGCTGCACTGTCCGACAGAGGCTGGAAGCAATAGCTCAAGCTTCTCTTC 1326
QY 1201 TCTCTGATATCAATCAATGTTTCAATTAATCAGCCAGCATGCGAGCATCTTACATGCTTC 1260
DB 1327 TCTCTGATATCAATCAATGTTTCAATTAATCAGCCAGCATGCGAGCATCTTACATGCTTC 1386
QY 1261 TCTCTATAGAAAGATGCTTTGGAATTAATCAAACTTCCACTACTCTGGAATGGAGCAAC 1320
DB 1387 TCTCTATAGAAAGATGCTTTGGAATTAATCAAACTTCCACTACTCTGGAATGGAGCAAC 1446
QY 1321 AAGCTATGCCAGTCTCCCTGTTTCAAGAACTATCCGAGAGAGATCCCGAAACAGTCTT 1380
DB 1447 AAGCTATGCCAGTCTCCCTGTTTCAAGAACTATCCGAGAGAGATCCCGAAACAGTCTT 1506
QY 1381 GATTAAGAGGAAGGACACATCTCCCAAGAGCTGACGCGCAGGCTTTCAGACAGCAG 1440
DB 1507 GATTAAGAGGAAGGACACATCTCCCAAGAGCTGACGCGCAGGCTTTCAGACAGCAG 1566
QY 1441 AGCAAGGATTTGCTGATGCTGTCAGAACAGAGAGAGTGGACCGCCAGAGTCCCTTTTA 1500
DB 1567 AGCAAGGATTTGCTGATGCTGTCAGAACAGAGAGAGTGGACCGCCAGAGTCCCTTTTA 1626
QY 1501 TCTTCACTGATGAGAGTGGAGCTGGAAGCAATTAACACACAGCTTCTTTTCGCG 1560
DB 1627 TCTTCACTGATGAGAGTGGAGCTGGAAGCAATTAACACACAGCTTCTTTTCGCG 1686
QY 1561 CTTTCCACAGCAGAGAGACCTTCAAGAGTGTGCTGGCTTGAAGGCTGGCAC 1620
DB 1687 CTTTCCACAGCAGAGAGACCTTCAAGAGTGTGCTGGCTTGAAGGCTGGCAC 1746
QY 1621 TCGGATATCTTGGCCCCCGACAGACTCTACCCCTTCCAGCAGAGCTGGATTTTGGC 1680
DB 1747 TCGGATATCTTGGCCCCCGACAGACTCTACCCCTTCCAGCAGAGCTGGATTTTGGC 1806
QY 1681 ACAGATCTCTACACTTCTACTCTGTGCTCAGCCATCTTACGAGAGGAGTCCAGTTACTCT 1740
DB 1807 ACAGATCTCTACACTTCTACTCTGTGCTCAGCCATCTTACGAGAGGAGTCCAGTTACTCT 1866
QY 1741 GCCTACAGCTGACAGCAGCTGCCCATTTGGAGAGCAAGTCTATTTCTGTGCGAGGCGG 1800
DB 1867 GCCTACAGCTGACAGCAGCTGCCCATTTGGAGAGCAAGTCTATTTCTGTGCGAGGCGG 1926
QY 1801 CAGAAAGCAAGTGAAGAGCTGACTCGCGCGAGCTGGCAATGAAGAGAGCCCTTTGAA 1860
DB 1927 CAGAAAGCAAGTGAAGAGCTGACTCGCGCGAGCTGGCAATGAAGAGAGCCCTTTGAA 1986
QY 1861 AAGCAGTTTAAACGAGAAAGCTGCCAAATGGAATTTTGAAGAGAGATCATGTCAAGAAAC 1920
DB 1987 AAGCAGTTTAAACGAGAAAGCTGCCAAATGGAATTTTGAAGAGAGATCATGTCAAGAAAC 2046
```

```

OY      1921 AGGTACCGGGAAGACTGGGGAAAGTGGGCAGTCAGCTTAGCTTTCCGGCAGCATGGAA 1980
Db      2047 AGGTACCGGGAAGACTGGGGAAAGTGGGCAGTCAGCTTAGCTTTCCGGCAGCATGGAA 2106

OY      1981 ATCATTTAGAGTCTCC 1995
        |||||
Db      2107 ATCATTTAGAGTCTCC 2121

RESULT 4
US-09-964-277-1
; Sequence 1, Application US/09964277
; Patent No. US2002013170A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434
; CURRENT APPLICATION NUMBER: US/09/964,277
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-277-1
```

Query Match	94.9%	Score 1893;	DB 9;	Length 3496;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1993; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

OY	1	ATGGCCCATGAGATGATATTGGAACTCAAAATTGTACTGAGAGGTTGGTGCCTGCGAA	60
Db	562	ATGGCCCATGAGATGATATTGGAACTCAAAATTGTACTGAGAGGTTGGTGCCTGCGAA	621
OY	61	AGTGAACGGAAGAAAGTGCCTAATTGATAGCGCGCAATTTGTGGATACAAATACATCC	120
Db	622	AGTGAACGGAAGAAAGTGCCTAATTGATAGCGCGCAATTTGTGGATACAAATACATCC	681
OY	121	CACATTTTGGAAAGCATTTAATATCAACCTGCTCCAAAGCTTATGAAAGCAAGTTGCACAG	180
Db	682	CACATTTTGGAAAGCATTTAATATCAACCTGCTCCAAAGCTTATGAAAGCAAGTTGCACAG	741
OY	181	GACAAAGTGTAAATTAACAGAGCTCATCCAGCAATTCAGGAAACATTAAGGTGACATTGAT	240
Db	742	GACAAAGTGTAAATTAACAGAGCTCATCCAGCAATTCAGGAAACATTAAGGTGACATTGAT	801
OY	241	TGCAGTCGAAGGTTGTATGATTACGATCAAAAGCTCCAAAGATGTGCTCTCTCTTCA	300
Db	802	TGCAGTCGAAGGTTGTATGATTACGATCAAAAGCTCCAAAGATGTGCTCTCTCTTCA	861
OY	301	GACTGTTTTCTCACTGTACTTCTGGGTAACTGGAAGAAAGCTTCACTGTTACAGTG	360
Db	862	GACTGTTTTCTCACTGTACTTCTGGGTAACTGGAAGAAAGCTTCACTGTTACAGTG	921
OY	361	CTTGAGAGTGGGTTGTGCTGAGTCTCTCGTGTGTTCCCTGGCCTCTGGAAGAAATCC	420
Db	922	CTTGAGAGTGGGTTGTGCTGAGTCTCTCGTGTGTTCCCTGGCCTCTGGAAGAAATCC	981
OY	421	ACTCTAGTCCCTACCTGCAATTTCTCAAGCTTGCCTTACTGTGTGCCAACTTTGGGCCAAC	480
Db	982	ACTCTAGTCCCTACCTGCAATTTCTCAAGCTTGCCTTACTGTGTGCCAACTTTGGGCCAAC	1041
OY	481	CGAATTTCTTCCCAATCTTTATCTTGGGCTGCCAGAGATGTTCCTCAACAAAGAGCTGATA	540
Db	1042	CGAATTTCTTCCCAATCTTTATCTTGGGCTGCCAGAGATGTTCCTCAACAAAGAGCTGATA	1101
OY	541	CAGCAGATGGGAATGGTTATGTGTTAATGCAAGCTTATCCTGCTCCAAAGCCTGACCTTT	600
Db	1102	CAGCAGATGGGAATGGTTATGTGTTAATGCAAGCTTATCCTGCTCCAAAGCCTGACCTTT	1161
OY	601	ATCCCGAGTCTCATTTCTCGCTGCTGCTGTGAATGACAGCTTTTGTGACAAATTTTG	660

Db	1162	ATCCCCAGTCTCATTTCTCGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAATTTTGG	1221
Oy	661	CCGTGTTTGGCAAAATCAGTAGATTTGATGAGAAAGCAAAGCCCTCCAAATGATGTGTT	720
Db	1222	CCGTGTTTGGCAAAATCAGTAGATTTTCAATTGAGAAAGCAAAGCCCTCCAAATGATGTGTT	1281
Oy	721	CTAGTGCATGTATTAGCTGGGATATCCCCGTCTCCGCCACATCCGCTATTCGCTACATCATG	780
Db	1282	CTAGTGCATGTATTAGCTGGGATATCCCCGTCTCCGCCACATCCGCTATTCGCTACATCATG	1341
Oy	781	AAGAGGATGACATGTCTTTAGATGAAGCTTACAGATTTTGAAAGAAAAAGACCTACT	840
Db	1342	AAGAGGATGACATGTCTTTAGATGAAGCTTACAGATTTTGAAAGAAAAAGACCTACT	1401
Oy	841	ATATCTCCAACTTCAAATTTTCTGGGGCAACTCTCTGACCTATGAGAAAGATTTAAGAAC	900
Db	1402	ATATCTCCAACTTCAAATTTTCTGGGGCAACTCTCTGACCTATGAGAAAGATTTAAGAAC	1461
Oy	901	CAGACTGGAGCATCAGGGCCCAAGAGCAAACTCAGCTGTGCACCTTGAGAGCCCAAT	960
Db	1462	CAGACTGGAGCATCAGGGCCCAAGAGCAAACTCAGCTGTGCACCTTGAGAGCCCAAT	1521
Oy	961	GAACCTGTCCCTGTGTCTCAGAGGGTGGACAGAAAAAGCGAGGCCCTCAGTCCACC	1020
Db	1522	GAACCTGTCCCTGTGTCTCAGAGGGTGGACAGAAAAAGCGAGGCCCTCAGTCCACC	1581
Oy	1021	TGTCCCACTCTGTGTAACCTCAGAGGCGACAGCAAAAGCCCTGTCAATCCCGCAGCGTG	1080
Db	1582	TGTCCCACTCTGTGTAACCTCAGAGGCGACAGCAAAAGCCCTGTCAATCCCGCAGCGTG	1641
Oy	1081	CCNAGCGTGCCAGCGTGACAGCCGTGTCTTTAGAGGACAGCCCGCTGTATCAGGCGCTC	1140
Db	1642	CCNAGCGTGCCAGCGTGACAGCCGTGTCTTTAGAGGACAGCCCGCTGTATCAGGCGCTC	1701
Oy	1141	AGTGGCTGTGACCTGTCCGACAGAGCGCTGGAAACAGCAATTAAGCTTCAAGCTTCTTC	1200
Db	1702	AGTGGCTGTGACCTGTCCGACAGAGCGCTGGAAACAGCAATTAAGCTTCAAGCTTCTTC	1761
Oy	1201	TCTCTGATATCAAAATCAGTTTCATATTTCAGCCAGCATGGCAGATCTTTCATGTGCTTC	1260
Db	1762	TCTCTGATATCAAAATCAGTTTCATATTTCAGCCAGCATGGCAGATCTTTCATGTGCTTC	1821
Oy	1261	TCCTCATCAGAAAGTCTTTGGAATTAACAACCTTCCACTACTCTGTGATGGGACCAAC	1320
Db	1822	TCCTCATCAGAAAGTCTTTGGAATTAACAACCTTCCACTACTCTGTGATGGGACCAAC	1881
Oy	1321	AAGCTATGCGAGTTCTCCCTGTTCAGAACTATCGGAGCAGACTCCCGAAACCAAGTCTC	1380
Db	1882	AAGCTATGCGAGTTCTCCCTGTTCAGAACTATCGGAGCAGACTCCCGAAACCAAGTCTC	1941
Oy	1381	GATTAAGAGGAAAGCAGACATCCGCCAAGCTGACACCGCAGGCTTTCAGACGCGAG	1440
Db	1942	GATTAAGAGGAAAGCAGACATCCGCCAAGCTGACACCGCAGGCTTTCAGACGCGAG	2001
Oy	1441	AGCAAGCGATTGCAATTCGATCAGAACCAAGCAGCAGTGGCAACCGCCAGAGGTCCCTTTTA	1500
Db	2002	AGCAAGCGATTGCAATTCGATCAGAACCAAGCAGCAGTGGCAACCGCCAGAGGTCCCTTTTA	2061
Oy	1501	TCTTCACTGCATCGAAGTGGGAGCGTGGAGGACAAATTACACACACAGCTTCTTTTGGGC	1560
Db	2062	TCTTCACTGCATCGAAGTGGGAGCGTGGAGGACAAATTACACACACAGCTTCTTTTGGGC	2121
Oy	1561	CTTTTCCACGACGACGACACCTCAGAGAGCTGTGGCCCTGGGCTTTAAGGGCTGGCAC	1620
Db	2122	CTTTTCCACGACGACGACACCTCAGAGAGCTGTGGCCCTGGGCTTTAAGGGCTGGCAC	2181
Oy	1621	TCCGATATCTTGGCCCCCAGACCTCTAACCCCTTTCCTGACAGACAGCTGTATTTTGGC	1680
Db	2182	TCCGATATCTTGGCCCCCAGACCTCTAACCCCTTTCCTGACAGACAGCTGTATTTTGGC	2241
Oy	1681	ACAGAGTCTTCAACTTCTACTGTGCTCAGCCACTTACGGAGGCAGTGCCAGTTACTCT	1740

Db	2242	ACAGAGTCTCAACACTTCTACTCTGCTCAGCCATCTACGGAGGACAGTCCAGTTACTCT	2301
QY	1741	GCTTACAGCTGCAGCCAGCTGGCCACTTGGCGAGACCAAGTCTATTCTGTGGCAGGCGG	1800
Db	2302	GCTTCAGCTGCAGCCAGCTGGCCACTTGGCGAGACCAAGTCTATTCTGTGGCAGGCGG	2361
QY	1801	CAGAGCCAAAGTGAAGAGCTGACTTCCTCGCGGCGAGCTGGCATGAAAGAGAGCCCTTTGAA	1860
Db	2362	CAGAGCCAAAGTGAAGAGCTGACTTCGCGGCGAGCTGGCATGAAAGAGAGCCCTTTGAA	2421
QY	1861	AAGCAGTTTAAACCGAGAAGCTGCCAAATTTGGAGAGAGCATCATGAGAGAC	1920
Db	2422	AAGCAGTTTAAACCGAGAAGCTGCCAAATTTGGAGAGAGCATCATGAGAGAC	2481
QY	1921	AGGTCAACGGGAAGAGCTGGGGGAAATGGGCGACTCAGTCTAGCTTTTGGGCGAGATGAA	1980
Db	2482	AGGTCAACGGGAAGAGCTGGGGGAAATGGGCGAGTCACTTCTTTGGGCGAGATGAA	2541
QY	1981	ATCATTTGAGGTCTCC	1995
Db	2542	ATCATTTGAGGTCTCC	2556

```

RESULT 5
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-03002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589) ... (2583)
; US-09-816-494-1

```

Query Match	94.9%	Score 1893;	DB 9;	Length 3544;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1993;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0

QY	1	ATGGCCCATGATGATTTGMACTCAAAATTTGTAAGAGGGTGTGGCTCTGCTGGAA	60
Db	589	ATGGCCCATGATGATTTGMACTCAAAATTTGTAAGAGGGTGTGGCTCTGCTGGAA	648
QY	61	AGTGAAGGAAAAAGTCTGCTTAATTGATAGCCGGCCATTTGTGAAATCAATATACATCC	120
Db	649	AGTGAAGGAAAAAGTCTGCTTAATTGATAGCCGGCCATTTGTGAAATCAATATACATCC	708
QY	121	CACATTTTGGAGCCATTAATATCAACTCTCCAGCTTATGAAGCGAAGGTGTCCAAAG	180
Db	709	CACATTTTGGAGCCATTAATATCAACTCTCCAGCTTATGAAGCGAAGGTGTCCAAAG	768
QY	181	GACAAAGTGTATTTCAGAGGCTCATCCGACATTCAGGAAACATTAAGTTGACATTTGAT	240
Db	769	GACAAAGTGTATTTCAGAGGCTCATCCGACATTCAGGAAACATTAAGTTGACATTTGAT	828
QY	241	TGCAGTCAGAGGTTTACGTTACGATCAAACTCCCAAGATGTGCTCTCTCTTCA	300
Db	829	TGCAGTCAGAGGTTTACGTTACGATCAAACTCCCAAGATGTGCTCTCTCTTCA	888
QY	301	GACTGTTTCTCACTGTACTTCTGGTAACTGGAGAAAGCTTCAACTCTGTTCACTTG	360

Db	889	GACTGTTTTCTCACTGTAACTTCTGGTAAACTGGAGAAGACTTCAACTCTGTTCACTG	948
QY	361	CTTGCAAGTGGGTTTGCTGAATTCCTCGTGTGTTTTCCCTGGACCTCTGTGAGAGAAATCC	420
Db	949	CTTGCAAGTGGGTTTGCTGAATTCCTCGTGTGTTTTCCCTGGACCTCTGTGAGAGAAATCC	100
QY	421	ACTCAAGTCCCTACCTCAATTTCTCAAGCCCTTGGCTTAACCTGTGCGCAACATTTGGGCCAAC	480
Db	1009	ACTCAAGTCCCTACCTCAATTTCTCAAGCCCTTGGCTTAACCTGTGCGCAACATTTGGGCCAAC	106
QY	481	CGAATCTTCCCAATCTTTATCTTGGCTGCAAGCATGTCTCTCAACAAGAGCTGATA	540
Db	1069	CGAATCTTCCCAATCTTTATCTTGGCTGCAAGCATGTCTCTCAACAAGAGCTGATA	112
QY	541	CAGAGAAATGGGANTGGTTATGTGTTAAATGCCAGCTAATCTGTCCAAAGCTGACTTT	600
Db	1129	CAGAGAAATGGGANTGGTTATGTGTTAAATGCCAGCAATACTGTCCAAAGCTGACTTT	1188
QY	601	ATCCCCAGATCAATTTCTCGCTGTGCTCGTAATGACAGCTTTTGTGAGAAATTTTG	660
Db	1189	ATCCCCAGATCAATTTCTCGCTGTGCTCGTAATGACAGCTTTTGTGAGAAATTTTG	1248
QY	661	CCGTGTGTGGACAATACAGTAATTTCAATTGAGAAAGCAAAAGCCTCCAAATGATGTGT	720
Db	1249	CCGTGTGTGGACAATACAGTAATTTCAATTGAGAAAGCAAAAGCCTCCAAATGATGTGT	1308
QY	721	CTATGCACCTGTTTACCTGGAGATCTCCCGCTCCGSCACCATGCTAATGGCCTACATCATG	780
Db	1309	CTATGCACCTGTTTACCTGGAGATCTCCCGCTCCGSCACCATGCTAATGGCCTACATCATG	1368
QY	781	AAGAAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGAAAGACTACT	840
Db	1369	AAGAAGATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGAAAGACTACT	1428
QY	841	ATATCTCCAAACTTCAATTTCTGGGCCAATCTCCCTGACCTATGAGAGAAAGATTTAAGAC	900
Db	1429	ATATCTCCAAACTTCAATTTCTGGGCCAATCTCCCTGACCTATGAGAGAAAGATTTAAGAC	1488
QY	901	CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACCTGAGAACCCAAAT	960
Db	1489	CAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACCTGAGAACCCAAAT	1548
QY	961	GAAACCTGTCCCGCTGTCTCAGAGGGGTGACAGAAAAGCGACGCCCCCTTCAATGCCACC	1020
Db	1549	GAAACCTGTCCCGCTGTCTCAGAGGGGTGACAGAAAAGCGACGCCCCCTTCAATGCCACC	1608
QY	1021	TGTGCCAATCTGTCACTCAGAGGACAGAGCAAAAGCCCCGTGATCCCGCAGGCTG	1080
Db	1609	TGTGCCAATCTGTCACTCAGAGGACAGAGCAAAAGCCCCGTGATCCCGCAGGCTG	1668
QY	1081	CCGACGCTGCCAAGCTGTGCAAGCCGTGCTTTAGAGGACAGCCCGCTGTGATCAGGGCTTC	1140
Db	1669	CCGACGCTGCCAAGCTGTGCAAGCCGTGCTTTAGAGGACAGCCCGCTGTGATCAGGGCTTC	1728
QY	1141	AGTGGGCTGCACCTGTCCGACAGACAGGCTGGAAGACAGAAATTAAGCTCAAGGCTTCTTC	1200
Db	1729	AGTGGGCTGCACCTGTCCGACAGACAGGCTGGAAGACAGAAATTAAGCTCAAGGCTTCTTC	1788
QY	1201	TCTCTGATCAAAATCAGTTTCAATTAATGACCGACAGATGAGACATCTTACATGAGCTTC	1260
Db	1789	TCTCTGATCAAAATCAGTTTCAATTAATGACCGACAGATGAGACATCTTACATGAGCTTC	1848
QY	1261	TCTCTCATCAGAAATGCTTTGGAAATCTAACAACTTTCCACTATCTTGAATGGAGCAAC	1320
Db	1849	TCTCTCATCAGAAATGCTTTGGAAATCTAACAACTTTCCACTATCTTGAATGGAGCAAC	1908
QY	1321	AAGCTATGCAGTTCTCCCTGTTGAGAACTATGAGACAGACTCCGAAACCAAGTCTT	1380
Db	1909	AAGCTATGCAGTTCTCCCTGTTGAGAACTATGAGACAGACTCCGAAACCAAGTCTT	1968
QY	1381	GATTAAGAGAAACCGACATCCCCAGAAAGCTGACAGATCCGCAAGGCTTTACAGACGCAAG	1440
Db	1969	GATTAAGAGAAACCGACATCCCCAGAAAGCTGACAGATCCGCAAGGCTTTACAGACGCAAG	2028

OY	1441	AGCAAGCCATTTCATTCGGCTCGAAGAACACACAGAGTGGACCGCCACAGAGTCCCTTTTA	1500
Db	2029	AGCAAGCCATTTCATTCGGCTCGAAGAACACACAGAGTGGACCGCCACAGAGTCCCTTTTA	2088
OY	1501	TCTCCACTGCATCGAAGTGGAGCGTGGAGGACAAATTACACACAGCTTCTTTTCGGC	1560
Db	2089	TCTCCACTGCATCGAAGTGGAGCGGTGGAGGACAAATTACACACAGCTTCTTTTCGGC	2148
OY	1561	CTTTTCACGACGACGACGACGACCTACACAAAGTCTGCTGGCTTGGCTTTAAGGCTGGCAC	1620
Db	2149	CTTTTCACGACGACGACGACGACCTACACAAAGTCTGCTGGCTTGGCTTTAAGGCTGGCAC	2208
OY	1621	TCGATATCTTGGGCCCCCAGACCTCTACCCCTCCCTGACCCAGACAGCTGATATTTCGCC	1680
Db	2209	TCGATATCTTGGGCCCCCAGACCTCTACCCCTCCCTGACCCAGACAGCTGATATTTCGCC	2268
OY	1681	ACAGAGTCTCTACACTTCTACTCTGTGCTGAGCCATCTACGAGGCGAGTCCAGTTACTCT	1740
Db	2269	ACAGAGTCTCTACACTTCTACTCTGTGCTGAGCCATCTACGAGGCGAGTCCAGTTACTCT	2328
OY	1741	GCCTACAGCTGCAGCCAGCTGCCCACTTCGGAGACCAAGTCTATTCTGTGGCAGGCGG	1800
Db	2329	GCCTACAGCTGCAGCCAGCTGCCCACTTCGGAGACCAAGTCTATTCTGTGGCAGGCGG	2388
OY	1801	CAGAGCCAGAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAGAGAGAGCCCTTTGAA	1860
Db	2389	CAGAGCCAGAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAGAGAGAGCCCTTTGAA	2448
OY	1861	AAGCAGTTTAAACGCGAAGAGCTGCGCAATGGAATTTGGAGAGCATCATGTTCAGAAAC	1920
Db	2449	AAGCAGTTTAAACGCGAAGAGCTGCGCAATGGAATTTGGAGAGCATCATGTTCAGAAAC	2508
OY	1921	AGGTACCGGGAAGAGCTGGGGGAAAGTGGGCACTCAGTCTTTCGGGCGAGCATGGA	1980
Db	2509	AGGTACCGGGAAGAGCTGGGGGAAAGTGGGCACTCAGTCTTTCGGGCGAGCATGGA	2568
OY	1981	ATCATTGAGGTCTCC 1995	
Db	2569	ATCATTGAGGTCTCC 2583	
RESULT 6			
US-10-377-072-25			
Sequence 25, Application US/10377072			
Publication No. US20040009501A1			
GENERAL INFORMATION:			
APPLICANT: Millennium Pharmaceuticals Inc.			
APPLICANT: Curtis, Rory A.J.			
APPLICANT: Logan, Thomas Joseph			
APPLICANT: Glucksmann, Maria A.			
APPLICANT: Meyers, Rachel E. J.			
APPLICANT: Williamson, Mark J.			
APPLICANT: Rudolph-Owen, Laura A.			
APPLICANT: Tsai, Feng-Ying			
APPLICANT: Chun, Miyoung			
TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,			
TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES			
TITLE OF INVENTION: AND USES THEREFOR			
FILE REFERENCE: MP103-0180NM1M			
CURRENT APPLICATION NUMBER: US/10/377, 072			
CURRENT FILING DATE: 2003-02-27			
PRIOR APPLICATION NUMBER: US 09/895, 860			
PRIOR FILING DATE: 2001-06-29			
PRIOR APPLICATION NUMBER: US 60/215, 370			
PRIOR FILING DATE: 2000-06-29			
PRIOR APPLICATION NUMBER: US 09/723, 806			
PRIOR FILING DATE: 2000-11-28			
PRIOR APPLICATION NUMBER: US 60/187, 455			
PRIOR FILING DATE: 2000-03-07			
PRIOR APPLICATION NUMBER: US 09/843, 297			
PRIOR FILING DATE: 2001-04-25			
PRIOR APPLICATION NUMBER: US 60/199, 801			

```

? PRIOR FILING DATE: 2000-04-26
? PRIOR APPLICATION NUMBER: US 09/861,801
? PRIOR FILING DATE: 2001-05-21
? PRIOR APPLICATION NUMBER: US 60/205,508
? PRIOR FILING DATE: 2000-05-19
? PRIOR APPLICATION NUMBER: US 09/816,494
? PRIOR FILING DATE: 2001-03-23
? PRIOR APPLICATION NUMBER: US 09/815,419
? PRIOR FILING DATE: 2001-03-22
? Remaining Prior Application data removed - See file wrapper or PALM.
? NUMBER OF SEQ ID NOS: 114
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 25
? LENGTH: 3544
? TYPE: DNA
? ORGANISM: Homo Sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (589) ... (2586)
US-10-377-072-25

```

Query Match	94.9%	Score 1893;	DB 16;	Length 3544;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1993; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	1	ATGGCCCATGAGATGATTTGAACTCAATATTTTACTGAGAGTGTGGCTCTGCTGGA	60
Db	589	ATGGCCCATGAGATGATTTGAACTCAATATTTTACTGAGAGTGTGGCTCTGCTGGA	648
QY	61	AGTGGAA CGGAAAAAGTCTGCTATTTAGTACCGGCGCATTTGTGGAAATACATACATCC	120
Db	649	AGTGGAA CGGAAAAAGTCTGCTATTTAGTACCGGCGCATTTGTGGAAATACATACATCC	708
QY	121	CACATTTTGGAAAGCATTAATATCACTGCTCCACGTTATGAAGCGCAAGGTTCACAAG	180
Db	709	CACATTTTGGAAAGCATTAATATCACTGCTCCACGTTATGAAGCGCAAGGTTCACAAG	768
QY	181	GACAAAGTGTAAATTACAGAGCTATCAGACATTCAGCGAAACATTAAGTTGACATTGAT	240
Db	769	GACAAAGTGTAAATTACAGAGCTATCAGACATTCAGCGAAACATTAAGTTGACATTGAT	828
QY	241	TGCAGTCAAGAGGTTTGATTTACGATCAAAAGCTCCCAAGTGTGGCTCTCTCTTCA	300
Db	829	TGCAGTCAAGAGGTTTGATTTAGATCAAAAGCTCCCAAGTGTGGCTCTCTCTTCA	888
QY	301	GACGTGTTTTCACCTGACTCTGAGGTAACTGAGAGAAGACCTTCAACTCTGTTCACTG	360
Db	889	GACGTGTTTTCACCTGACTCTGAGGTAACTGAGAGAAGACCTTCAACTCTGTTCACTG	948
QY	361	CTTGCAAGTGGGTTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGGAAGAAATCC	420
Db	949	CTTGCAAGTGGGTTTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGGAAGAAATCC	1008
QY	421	ACTTAGTCCCTACTCCTGATTTCTCAGCCTTGCTTAACCTGTGGCAATATGGGGCAAC	480
Db	1009	ACTTAGTCCCTACTCCTGATTTCTCAGCCTTGCTTAACCTGTGGCAATATGGGGCAAC	1068
QY	481	CGAATTTCTCCCAATCTTTATCTTGAGCTCCAGCGAGATGTCTCAACAGAGACTGATA	540
Db	1069	CGAATTTCTCCCAATCTTTATCTTGAGCTCCAGCGAGATGTCTCAACAGAGACTGATA	1128
QY	541	CAGCGAAATGGGATTTGGTTATGTGTAAATGCCAGCTATACCTGTCCAAACCTGACTTT	600
Db	1129	CAGCGAAATGGGATTTGGTTATGTGTAAATGCCAGCAATACCTGTCCAAACCTGACTTT	1188
QY	601	ATCCCGCAGTCTCATTTCTGCGCTGCTGCTGATGACAGCTTTTGTGAGAAAATTTTG	660
Db	1189	ATCCCGCAGTCTCATTTCTGCGCTGCTGCTGATGACAGCTTTTGTGAGAAAATTTTG	1248
QY	661	CCGTGTTTGGACAATCAGTAGATTTCAATTGAGAAAGCAAAAGCTCCATAGATGTGTT	720
Db	1249	CCGTGTTTGGACAATCAGTAGATTTCAATTGAGAAAGCAAAAGCTCCATAGATGTGTT	1308

QY 721 CTAGTCACGTTGTTAGCTGGATCTCCGGCTCCGCCACACATGGTATGGCTTACATCATG 780
Db 1309 CTAGTCACGTTGTTAGCTGGATCTCCGGCTCCGCCACATGGCTTACATCATG 1368
QY 781 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTTACT 840
Db 1369 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACTTACT 1428
QY 841 ATATCTCCAACTTCAATTTTCTGGGCCAATCCCTGACATATGAGAGAAATTAAGAAC 900
Db 1429 ATATCTCCAACTTCAATTTTCTGGGCCAATCCCTGACATATGAGAGAAATTAAGAAC 1488
QY 901 CAGACTGAGACATCAGGGCCAAAGACAACTCAGCTGTCACCTTGAGAACCAAT 960
Db 1489 CAGACTGAGACATCAGGGCCAAAGACAACTCAGCTGTCACCTTGAGAACCAAT 1548
QY 961 GAACCTGTCCTGCTGCTTCAAGAGGTGACAGAAAAAGGAGCGCCCTTCACTCACCC 1020
Db 1549 GAACCTGTCCTGCTGCTTCAAGAGGTGACAGAAAAAGGAGCGCCCTTCACTCACCC 1608
QY 1021 TGTGCGCACTGTGCTACTCAGAGGACAGACAGAAAGGCGCGTGCATCCCGCAGCGTG 1080
Db 1609 TGTGCGCACTGTGCTACTCAGAGGACAGACAGAAAGGCGCGTGCATCCCGCAGCGTG 1668
QY 1081 CCCAGCGTCCCAAGCGTGCAGCGCTCGCTGTTAGAGACAGCCCGCTGTATCAGAGCGCTG 1140
Db 1669 CCCAGCGTCCCAAGCGTGCAGCGCTCGCTGTTAGAGACAGCCCGCTGTATCAGAGCGCTG 1728
QY 1141 AGTGGGCTGCACCTGTCCGACAGACAGGCTGGAGAGACAGAAATTAAGCTCAAGCTTCTTC 1200
Db 1729 AGTGGGCTGCACCTGTCCGACAGACAGGCTGGAGAGACAGAAATTAAGCTCAAGCTTCTTC 1788
QY 1201 TCTCTGATATCAATTAAGTTTCAATTAATTCAGCCAGATGCGACATCTTACATGCTTC 1260
Db 1789 TCTCTGATATCAATTAAGTTTCAATTAATTCAGCCAGATGCGACATCTTACATGCTTC 1848
QY 1261 TCCCTATCAGAAAGTGTGTTGAAATCTAACAACCTTCCACTCTGTGATGGAGCAAC 1320
Db 1849 TCCCTATCAGAAAGTGTGTTGAAATCTAACAACCTTCCACTCTGTGATGGAGCAAC 1908
QY 1321 AAGCTATGCCAGTTCTCCCTGTTTCAGAACTATCGGAGACAGCTCCGAAACCAAGTCT 1380
Db 1909 AAGCTATGCCAGTTCTCCCTGTTTCAGAACTATCGGAGACAGCTCCGAAACCAAGTCT 1968
QY 1381 GATTAAGAGAGAAACCCAGCATCCCAAGAAAGCTGCAGACCGCGCTTACAGACAGCAG 1440
Db 1969 GATTAAGAGAGAAACCCAGCATCCCAAGAAAGCTGCAGACCGCGCTTACAGACAGCAG 2028
QY 1441 AGCAAGCGATTGCTATCGGTCAAGAACAGACAGAGTGGACCGCCAGAGGTCCCTTTTA 1500
Db 2029 AGCAAGCGATTGCTATCGGTCAAGAACAGACAGAGTGGACCGCCAGAGGTCCCTTTTA 2088
QY 1501 TCTCCACTGCTGATGAAGTGGAGAGCGTGAAGACAAATTACACACAGGCTCTTTTGGGC 1560
Db 2089 TCTCCACTGCTGATGAAGTGGAGAGCGTGAAGACAAATTACACACAGGCTCTTTTGGGC 2148
QY 1561 CTTTCCACAGCAGAGACAGCACTCAAGAAAGTGTGCTGCGCTTAAAGGCTGCGAC 1620
Db 2149 CTTTCCACAGCAGAGACAGCACTCAAGAAAGTGTGCTGCGCTTAAAGGCTGCGAC 2208
QY 1621 TGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGAGCTGTATTTTGGC 1680
Db 2209 TGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGAGCTGTATTTTGGC 2268
QY 1681 ACAGAGTCTCACTATTTCTATCTTGGCTGACAGCATCTACGAGGCAAGTCCAGTTACTCT 1740
Db 2269 ACAGAGTCTCACTATTTCTATCTTGGCTGACAGCATCTACGAGGCAAGTCCAGTTACTCT 2328
QY 1741 GCTTACAGCTGACAGCACTGCCACTTGGAGAGCAAGCTTATTTGTCGAGAGCGG 1800
Db 2329 GCTTACAGCTGACAGCACTGCCACTTGGAGAGCAAGCTTATTTGTCGAGAGCGG 2388
QY 1801 CAGAAAGCAAGTGAAGCTGACTCGCGGCGGAGCTGGACATGAAGAAGACCCCTTTGAA 1860

Db 2389 CAGAAAGCAAGTGAAGCTGACTCGCGCGGAGACTGGACATGAAGAAGACCCCTTTGAA 2448
QY 1861 AAGCACTTTTAAACGCAAGAGCTGCGCAATGAAATTTGAGAGAGCATGTCAGAGAAC 1920
Db 2449 AAGCACTTTTAAACGCAAGAGCTGCGCAATGAAATTTGAGAGAGCATGTCAGAGAAC 2508
QY 1921 AGTCCAGGGAAGAGCTGGGGAAGTGGGAGTCACTGCTAGCTTTTGGGCAAGCATGAA 1980
Db 2509 AGTCCAGGGAAGAGCTGGGGAAGTGGGAGTCACTGCTAGCTTTTGGGCAAGCATGAA 2568
QY 1981 ATCATTTAGAGTCTCC 1995
Db 2569 ATCATTTAGAGTCTCC 2583

RESULT 7
US-10-425-114-26234
; Sequence 26234, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 26234
; LENGTH: 3625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-028-HE_FLI
US-10-425-114-26234

Query Match 94.9%; Score 1893; DB 13; Length 3625;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCCATGATGATTTGGAATCAATTTGTTACTGAGAGTGGTGGCTGTGTA 60
Db 692 ATGGCCATGATGATTTGGAATCAATTTGTTACTGAGAGTGGTGGCTGTGTA 751
QY 61 AGTGAACGGAAAAGTGTGCTAATTTGATAGCCGCCATTTGTGAATCAATCATTC 120
Db 752 AGTGAACGGAAAAGTGTGCTAATTTGATAGCCGCCATTTGTGAATCAATCATTC 811
QY 121 CACATTTGGAAAGCCATTAATATCACTGCTCAAGCTTATGAACGAAAGTTGCAACG 180
Db 812 CACATTTGGAAAGCCATTAATATCACTGCTCAAGCTTATGAACGAAAGTTGCAACG 871
QY 181 GACAAAGTGTATTAACAGAGCTCATCCAGCTTACAGGAAACATTAAGTTGACATTGAT 240
Db 872 GACAAAGTGTATTAACAGAGCTCATCCAGCTTACAGGAAACATTAAGTTGACATTGAT 931
QY 241 TGCAGTCAGAAAGTGTATTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTTCA 300
Db 932 TGCAGTCAGAAAGTGTATTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTTCA 991
QY 301 GACTGTTTCTCATCTGTACTTGGTAACTGGAAGAGCTTCAACTCTGTTACCTG 360
Db 992 GACTGTTTCTCATCTGTACTTGGTAACTGGAAGAGCTTCAACTCTGTTACCTG 1051
QY 361 CTTCGAGGTGGTGTGCTGAGTCTCTGTTGTTTCCCTGGCCCTGTGAAAGAAATTC 420
Db 1052 CTTCGAGGTGGTGTGCTGAGTCTCTGTTGTTTCCCTGGCCCTGTGAAAGAAATTC 1111

QY	421	ACTCAGTCCCACTGGATTTCTCAGGCTGTCTTACCTGTGTGCAAACTTTGGGCGCAAC	480
Db	1112	ACTCAGTCCCACTGGATTTCTCAGGCTGTCTTACCTGTGTGCAAACTTTGGGCGCAAC	1171
QY	481	CGAATTTCTCCCAATCTTTATCTTGGGCGCGCAGCAGATGTCTCAACAAGACCTGATA	540
Db	1172	CGAATTTCTCCCAATCTTTATCTTGGGCGCGCAGCAGATGTCTCAACAAGACCTGATA	1231
QY	541	CAGCAGAAATGGGATTTGGTTATGTGTAAATGCGACGTATACCTGTCCAAAGCTGACTTT	600
Db	1232	CAGCAGAAATGGGATTTGGTTATGTGTAAATGCGACCAATATCCTGTCCAAAGCTGACTTT	1291
QY	601	ATCCCCGAGTCTCATTTCTCGTGCTGTGCTGTGAATGACAGCTTTTGTGAGAAAATTTTG	660
Db	1292	ATCCCCGAGTCTCATTTCTCGTGCTGTGCTGTGAATGACAGCTTTTGTGAGAAAATTTTG	1351
QY	661	CGGTGGTTGGACCAATCCAGTAGATTCTCATTTGGAAAGCAAAAGCTCCAAATGGATGTGTT	720
Db	1352	CGGTGGTTGGACCAATCCAGTAGATTCTCATTTGGAAAGCAAAAGCTCCAAATGGATGTGTT	1411
QY	721	CTAATGCACTGTTTATGCTGGGATCTCCGCTCCGCAACATCGCTATGCGCTGACATATG	780
Db	1412	CTAATGCACTGTTTATGCTGGGATCTCCGCTCCGCAACATCGCTATGCGCTGACATATG	1471
QY	781	AAGAGGATGGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAAGAAAAGACTACT	840
Db	1472	AAGAGGATGGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAAGAAAAGACTACT	1531
QY	841	ATATCTCCAAACTTCAATTTTCTGGGGCAACTCCCTGGACTTATGGAAGAGAATTAAGAC	900
Db	1532	ATATCTCCAAACTTCAATTTTCTGGGGCAACTCCCTGGACTTATGGAAGAGAATTAAGAC	1591
QY	901	CAGACTGAGGATCAGGGCGCAAAAGCAAACTCAAGCTGCTGCACTTGGAGAGCCAAAT	960
Db	1592	CAGACTGAGGATCAGGGCGCAAAAGCAAACTCAAGCTGCTGCACTTGGAGAGCCAAAT	1651
QY	961	GAACTGTCTCCGTCTGTCTCAGAGGGGTGACAGAAAACGAGAGCGCCCTCAGTCCACC	1020
Db	1652	GAACTGTCTCCGTCTGTCTCAGAGGGGTGAGCAAGAAAACGAGAGCGCCCTCAGTCCACC	1711
QY	1021	TGTGCGGACTCTGCTACCTCAGAGGCAACAAGAGCAAGAGCCCTGCACTCCGCGCAGGTG	1080
Db	1712	TGTGCGGACTCTGCTACCTCAGAGGCAACAAGAGCAAGAGCCCTGCACTCCGCGCAGGTG	1771
QY	1081	CCCAAGCTGCGCAGGCTGACGCGTGTGCTGTTAGAGACACACCGCTGTGACAGCGCTG	1140
Db	1772	CCCAAGCTGCGCAGGCTGACGCGTGTGCTGTTAGAGACACACCGCTGTGACAGCGCTG	1831
QY	1141	AGTGGGCTGACCTGTCTCCGCGACAGAGGTGGAAGACAGCAATTAAGCTTCAAGCTTCTTC	1200
Db	1832	AGTGGGCTGACCTGTCTCCGCGACAGAGGTGGAAGACAGCAATTAAGCTTCAAGCTTCTTC	1891
QY	1201	TCTCTGGAATTCGAATTCAGTTTCATATTTACGCGACATGGGACGATCTTACATGGCTTC	1260
Db	1892	TCTCTGGAATTCGAATTCAGTTTCATATTTACGCGACATGGGACGATCTTACATGGCTTC	1951
QY	1261	TCTCTATCAGAAAGTCTTTGGAACTACAAACCTTCACTAATCTGTGGATGGGACCAAC	1320
Db	1952	TCTCTATCAGAAAGTCTTTGGAACTACAAACCTTCACTAATCTGTGGATGGGACCAAC	2011
QY	1321	AAGCTATGCAAGTTCTCCCTGTGTCAGGAATATGAGAGAGACTCCCGAAACCAAGTCTT	1380
Db	2012	AAGCTATGCAAGTTCTCCCTGTGTCAGGAATATGAGAGAGACTCCCGAAACCAAGTCTT	2071
QY	1381	GATTAAGAGGAAGCCAGCATCCCCAAAGAGTGTGACACCGCCAGGCTTTGACAGCGAG	1440
Db	2072	GATTAAGAGGAAGCCAGCATCCCCAAAGAGTGTGACACCGCCAGGCTTTGACAGCGAG	2131
QY	1441	AGCAAGCAATTCATTTGGGTGCAAGACCAAGCAGCAGTGGCACCGCCCAAGAGTCCCTTTTA	1500
Db	2132	AGCAAGCAATTCATTTGGGTGCAAGACCAAGCAGCAGTGGCACCGCCCAAGAGTCCCTTTTA	2191
QY	1501	TCTCCACTGCAATGGAAGTGGGAGCGGTGAGGACAAATTAACAACCAAGCTTCTTTTGGCG	1560

```

Db      2192  TCTCCACTGCATCGAAGTGGAGCGCTGGAGAGCAATTACACACACACTTCTTTTGCC 2251
QY      1561  CTTTCCACCAAGCCAGACGACACTTCACCAAGTCTGCTGGCTGGCCCTTAAAGGCTGGCAC 1620
Db      2252  CTTTCCACCAAGCCAGACGACACTTCACCAAGTCTGCTGGCTGGCCCTTAAAGGCTGGCAC 2311
QY      1621  TCGATATCTTGGCCCCCAGACTCTTACCCCTTCCCTGACGACAGCTGATATTTGGC 1680
Db      2312  TCGATATCTTGGCCCCCAGACTCTTACCCCTTCCCTGACGACAGCTGATATTTGGC 2371
QY      1681  ACAGAGTCTCAGACTTCTACTCTGCGCTCAAGCCATCTACGGAAGCAGTGCAGTTACTCT 1740
Db      2372  ACAAGTCTCAGACTTCTACTCTGCGCTCAGCCATCTACGGAAGCAGTGCAGTTACTCT 2431
QY      1741  GCTTACAGCTGCAGACCAAGCTGCCACTTGGCGGAGACCAAGTCTATTTCTGTCGCGAGCG 1800
Db      2432  GCCTACAGCTGCAGACCAAGCTGCCACTTGGCGGAGACCAAGTCTATTTCTGTCGCGAGCG 2491
QY      1801  CAGAAGCCAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAA 1860
Db      2492  CAGAAGCCAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAA 2551
QY      1861  AAGCAGTTTAAAGCCAGAGCTGCCAATTTGGAATTTGGAAGAGACATCATGAGAGAAC 1920
Db      2552  AAGCAGTTTAAAGCCAGAGCTGCCAATTTGGAATTTGGAAGAGACATCATGAGAGAAC 2611
QY      1921  AGTTCAGGGGAAGCTGGGGAAAGTGGGAGTCACTGACTTTTGGGCGAGCATGAA 1980
Db      2612  AGTTCAGGGGAAGCTGGGGAAAGTGGGAGTCACTGACTTTTGGGCGAGCATGAA 2671
QY      1981  ATCATTGAGGTCTCC 1995
Db      2672  ATCATTGAGGTCTCC 2686

RESULT 8
US-10-343-357-17
; Sequence 17, Application US/10343357
; Publication No. US20040058341A1
GENERAL INFORMATION:
APPLICANT: INCTE GENOMICS, INC.; TANG, Y.Tom
APPLICANT: ELIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
APPLICANT: YAO, Monique G.; BURFORD, Neil
APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
APPLICANT: LEE, Ernestine A.; HARPLIN, April J.A.
APPLICANT: LU, Dyang Aina M.; TRIBOULEY, Catherine M.
APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
APPLICANT: YUE, Henry; WARREN, Bridget A.
APPLICANT: NGUYEN, Daniel B.; CHAWLA, Narinder K.
APPLICANT: KEARNEY, Liam
TITLE OF INVENTION: PROTEIN PHOSPHATASES
FILE REFERENCE: PI-0173 PCT
CURRENT APPLICATION NUMBER: US/10/343,357
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: PCT/US01/23716
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,679
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/223,272
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/224,309
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 60/226,728
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 60/229,254
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: US 60/231,366
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
; SEQ ID NO 17

```

LENGTH: 3766
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570C81
US-10-343-357-17

Query Match 94.9%; Score 1893; DB 13; Length 3766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATATTGAACTCAATTGTTACTGAGAGTGGTGGCTCTGCTGAA 60
Db 538 ATGGCCCATGAGATATTGAACTCAATTGTTACTGAGAGTGGTGGCTCTGCTGAA 597
QY 61 AGTGAACGGAAGAGTCTGCTAATTGATAGCCGGCCATTGTTGGAAATACATATACC 120
Db 598 AGTGAACGGAAGAGTCTGCTAATTGATAGCCGGCCATTGTTGGAAATACATATACC 657
QY 121 CACATTTTGAAGCATTATATATCACTGCTCAAGCTTATGAGCGAAGTTGCAACAG 180
Db 658 CACATTTTGAAGCATTATATATCACTGCTCAAGCTTATGAGCGAAGTTGCAACAG 717
QY 181 GACAAAGCTTAATTAACAGAGCTCATCCAGATTCAACGAAACATTAAGTTGACATTGAT 240
Db 718 GACAAAGCTTAATTAACAGAGCTCATCCAGATTCAACGAAACATTAAGTTGACATTGAT 777
QY 241 TGCAGTCAGAAAGTTGATGATTAAGATCAAAAGTCCCAAGATGTTGCTCTCTCTCA 300
Db 778 TGCAGTCAGAAAGTTGATGATTAAGATCAAAAGTCCCAAGATGTTGCTCTCTCTCA 837
QY 301 GACTGTTTCTCACTGTAATTTCTGGATTAACCTGAGAAAGCTTCAACTGTTTCACTG 360
Db 838 GACTGTTTCTCACTGTAATTTCTGGATTAACCTGAGAAAGCTTCAACTGTTTCACTG 897
QY 361 CTTCAGAGTGGGTTTGCTGAGTCTCTGCTGTTTCCCTGGCTCTGTGAAGAAATCC 420
Db 898 CTTCAGAGTGGGTTTGCTGAGTCTCTGCTGTTTCCCTGGCTCTGTGAAGAAATCC 957
QY 421 ACTCTAGTCCCTACCTGATCTTCTCAGCCTTGCTTACCTGTTGCCAACTTGGGCCAAC 480
Db 958 ACTCTAGTCCCTACCTGATCTTCTCAGCCTTGCTTACCTGTTGCCAACTTGGGCCAAC 1017
QY 481 CGAATTTCTCCCATCTTTATCTTGGCTGCGACGAGATGCTCTCAACAGAGCTGATA 540
Db 1018 CGAATTTCTCCCATCTTTATCTTGGCTGCGACGAGATGCTCTCAACAGAGCTGATA 1077
QY 541 CAGCAATATGGAGTGGTATGTTAATGCAAGCTATACCTGTCCAAAGCTTACTTT 600
Db 1078 CAGCAATATGGAGTGGTATGTTAATGCAAGCTATACCTGTCCAAAGCTTACTTT 1137
QY 601 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 660
Db 1138 ATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATGACAGCTTTTGTGAAGAAATTTTG 1197
QY 661 CCGTGGTTGGACAAATACATATGATTGATGGAAGCAAAAGCTTCAATGATGTT 720
Db 1198 CCGTGGTTGGACAAATACATATGATTGATGGAAGCAAAAGCTTCAATGATGTT 1257
QY 721 CTAGTGCACTGTTTATGCTGGAGTCTCCGCTCCGACACCATGCTATACATCATG 780
Db 1258 CTAGTGCACTGTTTATGCTGGAGTCTCCGCTCCGACACCATGCTATACATCATG 1317
QY 781 AAGAGATGAGCATGTTTATGATGAGCTTAACAGATTGTGAAGAAAGAAAGCTTACT 840
Db 1318 AAGAGATGAGCATGTTTATGATGAGCTTAACAGATTGTGAAGAAAGAAAGCTTACT 1377
QY 841 ATATCTCAAACTTCAATTTCTGGGCGCACTCTCGAGCTATGGAAGAGATTAAAGAC 900
Db 1378 ATATCTCAAACTTCAATTTCTGGGCGCACTCTCGAGCTATGGAAGAGATTAAAGAC 1437
QY 901 CAGACTGAGATCATGGGCCAAAGCAAACTCAAGCTGTGACCTGAGAAAGCAAAT 960

Db 1438 CAGACTGAGATCATGGGCCAAAGCAAACTCAAGCTGTGACCTGAGAAAGCAAAT 1497
QY 961 GAACCTGTCCCTGTCTCTCTCAAGAGGTGACAGAAAGCGAGAGCCCTCAGTCCACCC 1020
Db 1498 GAACCTGTCCCTGTCTCTCTCAAGAGGTGACAGAAAGCGAGAGCCCTCAGTCCACCC 1557
QY 1021 TGTGCGCATCTGCTACCTCAGAGCGACAGAGCAAAAGCCCGTCAATCCCGCCAGCTG 1080
Db 1558 TGTGCGCATCTGCTACCTCAGAGCGACAGAGCAAAAGCCCGTCAATCCCGCCAGCTG 1617
QY 1081 CCCAGGTCCTCAGAGGTGACGCTGCTGTTAGAGACAGCCCGCTGTACAGCGCTC 1140
Db 1618 CCCAGGTCCTCAGAGGTGACGCTGCTGTTAGAGACAGCCCGCTGTACAGCGCTC 1677
QY 1141 AGTGGGCTSCACTGTCGCGACAGAGGCTGGAAGACACCAATTAAGCTTAAGCTTCTTC 1200
Db 1678 AGTGGGCTSCACTGTCGCGACAGAGGCTGGAAGACACCAATTAAGCTTCTTC 1737
QY 1201 TCTTGATATCAAAATCAGTTTCATATTCAAGCAGATGAGCATCTTACATGGCTTC 1260
Db 1738 TCTTGATATCAAAATCAGTTTCATATTCAAGCAGATGAGCATCTTACATGGCTTC 1797
QY 1261 TCCCTATGAAGAGATGCTTTGGAATCTAACAACTTCCACTTCTGATGGAGCAAC 1320
Db 1798 TCCCTATGAAGAGATGCTTTGGAATCTAACAACTTCCACTTCTGATGGAGCAAC 1857
QY 1321 AAGCTATCCAGTTCCTCCCTGTTCAAGAACTATTCGAGAGAGATCCCGAAACCAAGTCT 1380
Db 1858 AAGCTATCCAGTTCCTCCCTGTTCAAGAACTATTCGAGAGAGATCCCGAAACCAAGTCT 1917
QY 1381 GATTAAGAGGAAGCAGCATTCCTCAAGAGCTGACAGCCGACGCTTCAACAGCAG 1440
Db 1918 GATTAAGAGGAAGCAGCATTCCTCAAGAGCTGACAGCCGACGCTTCAACAGCAG 1977
QY 1441 AGCAAGCGATGCAATTCGATCGACAGACAGAGAGTCCGCCCAAGAGTCCCTTTTA 1500
Db 1978 AGCAAGCGATGCAATTCGATCGACAGACAGAGAGTCCGCCCAAGAGTCCCTTTTA 2037
QY 1501 TCTCACTGCATCGAATGGAGAGGTGAGAGCAATTAACCAACAGCTTCTTTTCGGC 1560
Db 2038 TCTCACTGCATCGAATGGAGAGGTGAGAGCAATTAACCAACAGCTTCTTTTCGGC 2097
QY 1561 CTTTCCACAGCAGCAGACGCTTCAAGAGTCTGCTGGCTTGAAGGCTGGCAC 1620
Db 2098 CTTTCCACAGCAGCAGACGCTTCAAGAGTCTGCTGGCTTGAAGGCTGGCAC 2157
QY 1621 TCGGATATCTGGCCCCCAGACCTCTACCCCTTCCCTGACAGAGCTGGTATTTTGGC 1680
Db 2158 TCGGATATCTGGCCCCCAGACCTCTACCCCTTCCCTGACAGAGCTGGTATTTTGGC 2217
QY 1681 ACAGAGTCTCTCAACATTTCTGCTGCTCAAGCCATCTACGAGAGAGTCCAGTTACTCT 1740
Db 2218 ACAGAGTCTCTCAACATTTCTGCTGCTCAAGCCATCTACGAGAGAGTCCAGTTACTCT 2277
QY 1741 GCTTACAGCTGACGAGCGCCACTTGGCGAGACCAAGTCTTATTTCTGTGCGAGCGG 1800
Db 2278 GCTTACAGCTGACGAGCGCCACTTGGCGAGACCAAGTCTTATTTCTGTGCGAGCGG 2337
QY 1801 CAGAAACCAAGTGAACAGACTGACTCGCGCGAGCTGGCATGAAGAGCCCTTTTAA 1860
Db 2338 CAGAAACCAAGTGAACAGACTGACTCGCGCGAGCTGGCATGAAGAGCCCTTTTAA 2397
QY 1861 AAGCAGTTTAAACGAGAAAGCTTCCAAATGAAATTTGAGAGAGCATCATGTCAAGAAC 1920
Db 2398 AAGCAGTTTAAACGAGAAAGCTTCCAAATGAAATTTGAGAGAGCATCATGTCAAGAAC 2457
QY 1921 AAGTCAACGGAAGAGCTGGGAAAGTGGGCACTGCTACTTTTGGGAGAGATGGA 1980
Db 2458 AAGTCAACGGAAGAGCTGGGAAAGTGGGCACTGCTACTTTTGGGAGAGATGGA 2517
QY 1981 ATCATTTGAGGTCTCC 1995

Db 2518 ATCATTGAGGTCTCC 2532

RESULT 9
US-10-648-593-115
; Sequence 115, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 115
; LENGTH: 4790
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-648-593-115

Query Match 94.9%; Score 1893; DB 17; Length 4790;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCCATGAGATGATTGGAACCTCAATTGTTACTGAGAGGTTGGTGGCTCTGCTGGAA 60
Db 184 ATGGCCCATGAGATGATTGGAACCTCAATTGTTACTGAGAGGTTGGTGGCTCTGCTGGAA 243
QY 61 AGTGAACGGAAGAGTGTCTAATTGATAGCCGCGCATTTGTGAAATACATATACC 120
Db 244 AGTGAACGGAAGAGTGTCTAATTGATAGCCGCGCATTTGTGAAATACATATACC 303
QY 121 CACATTTTGAAGCCATTAATATCACTGCTCCAGTTTGAAGCGAAGTTGGCAACAG 180
Db 304 CACATTTTGAAGCCATTAATATCACTGCTCCAGTTTGAAGCGAAGTTGGCAACAG 363
QY 181 GACAAAGTGTATTAATACAGAGCTCATCCAGCATTCAGCGAACAATAGTTGACATTGAT 240
Db 364 GACAAAGTGTATTAATACAGAGCTCATCCAGCATTCAGCGAACAATAGTTGACATTGAT 423
QY 241 TGCAGTCAGAAAGTTGATGTTTACGATCAAAAGCTCCAGAGTGTGCTCTCTCTCA 300
Db 424 TGCAGTCAGAAAGTTGATGTTTACGATCAAAAGCTCCAGAGTGTGCTCTCTCTCA 483
QY 301 GACTGTTTCTCACTGACTCTTGCGGTAACTGGAGAAGCTTCAACTCTGTTCACTG 360
Db 484 GACTGTTTCTCACTGACTCTTGCGGTAACTGGAGAAGCTTCAACTCTGTTCACTG 543
QY 361 CTGGCAGGAGGTTTGTGAGTCTCTCGTGTGTTTCCCTGGCCCTGTGAAAGAAATCC 420
Db 544 CTGGCAGGAGGTTTGTGAGTCTCTCGTGTGTTTCCCTGGCCCTGTGAAAGAAATCC 603
QY 421 ACTCTAGTCCCTACCTGATCTTCTCAGCCTTGTCTTACCTGTGCAACATTTGGCCCAAC 480
Db 604 ACTCTAGTCCCTACCTGATCTTCTCAGCCTTGTCTTACCTGTGCAACATTTGGCCCAAC 663
QY 481 CGAATTTCTCCCAATCTTTATCTTGGCTGCGCAGCGAGATGCTCTCAACAAGAGCTGATG 540
Db 664 CGAATTTCTCCCAATCTTTATCTTGGCTGCGCAGCGAGATGCTCTCAACAAGAGCTGATG 723
QY 541 CAGAGAGATGGGATGTTGTTATGTTAAATGCGACGATATACCTGTCCAAAGCCGACTTT 660
Db 724 CAGAGAGATGGGATGTTGTTATGTTAAATGCGACGATATACCTGTCCAAAGCCGACTTT 783
QY 601 ATCCCGAGTCTCATTTCTGCGGTGCTGCTGTAATGACAGCTTTTGTGAAAAATTTTG 660
Db 784 ATCCCGAGTCTCATTTCTGCGGTGCTGCTGTAATGACAGCTTTTGTGAAAAATTTTG 843

QY 661 CCGTGGTTGGAACAATCAGATGATTTCAATGAGAAAGCAAAAGCTCCATGATGTTGT 720
Db 844 CCGTGGTTGGAACAATCAGATGATTTCAATGAGAAAGCAAAAGCTCCATGATGTTGT 903
QY 721 CTAGTCACTGTTTAACTGGAGATCTCCGCTCCGCCACCATCTGCTATTCGCTACATCATG 780
Db 904 CTAGTCACTGTTTAACTGGAGATCTCCGCTCCGCCACCATCTGCTATTCGCTACATCATG 963
QY 781 AAGAGATGACATGCTTTTAAATGAGCTTACAGATTTTGTGAAAAAAGAAAGAACTACT 840
Db 964 AAGAGATGACATGCTTTTAAATGAGCTTACAGATTTTGTGAAAAAAGAAAGAACTACT 1023
QY 841 ATATCTCCCAACTCAATTTTTCGGGCAACTCTCGGACATATGAGAAAGATTAAGAAC 900
Db 1024 ATATCTCCCAACTCAATTTTTCGGGCAACTCTCGGACATATGAGAAAGATTAAGAAC 1083
QY 901 CAGACTGAGCATCAGAGGCGCAAAAGCAAACTCAAGCTCTGCACTGAGAGCAACCAAT 960
Db 1084 CAGACTGAGCATCAGAGGCGCAAAAGCAAACTCAAGCTCTGCACTGAGAGCAACCAAT 1143
QY 961 GAACTGTCTCTGCTGTCTCAGAGGTTGACAGAAAAAGCGAGAGCCCTCACTCAACC 1020
Db 1144 GAACTGTCTCTGCTGTCTCAGAGGTTGACAGAAAAAGCGAGAGCCCTCACTCAACC 1203
QY 1021 TGTGCGGACTCTGCTACTCCTCAGAGGAGAGAGCAAAAGCCGTGATCCCGCAGCGTG 1080
Db 1204 TGTGCGGACTCTGCTACTCCTCAGAGGAGAGAGCAAAAGCCGTGATCCCGCAGCGTG 1263
QY 1081 CCCAGCGTCCCGCAGCGTGCAGCCGTGCTTGAAGAGAGCCGCTGCTGATCAGAGGCTC 1140
Db 1264 CCCAGCGTCCCGCAGCGTGCAGCCGTGCTTGAAGAGAGAGCCGCTGCTGATCAGAGGCTC 1323
QY 1141 AGTGGCTGCACTGTCTGCGCAGACAGGCTGGAAGACAGCAATAGCTCAAGCTCTTC 1200
Db 1324 AGTGGCTGCACTGTCTGCGCAGACAGGCTGGAAGACAGCAATAGCTCAAGCTCTTC 1383
QY 1201 TCTCTGATATCAAAACAGTTTCAATTCATTCAGCCAGATGCGAGATCTTACATGCTTC 1260
Db 1384 TCTCTGATATCAAAACAGTTTCAATTCATTCAGCCAGATGCGAGATCTTACATGCTTC 1443
QY 1261 TCCTCATCAGAAAGATCTTTGGAATCTACAAACCTTTCACTACTCTGATGGAAGCAAC 1320
Db 1444 TCCTCATCAGAAAGATCTTTGGAATCTACAAACCTTTCACTACTCTGATGGAAGCAAC 1503
QY 1321 AAGCTATGCGAAGTTCTCCCTGTTTCAGGAATATGAGAGCAGATCCCGAAACAGTCT 1380
Db 1504 AAGCTATGCGAAGTTCTCCCTGTTTCAGGAATATGAGAGCAGATCCCGAAACAGTCT 1563
QY 1381 GATTAAGAGAAAGCCAGCATTCGCCAAGAAAGCTGCAACCCGCGAGCTTCAAGCAGCCAG 1440
Db 1564 GATTAAGAGAAAGCCAGCATTCGCCAAGAAAGCTGCAACCCGCGAGCTTCAAGCAGCCAG 1623
QY 1441 AGCAAGCATTTGATTCGGTTCAGAAACCAAGCAGAGTGGCAGCCGCAAGAGTCCCTTTTA 1500
Db 1624 AGCAAGCATTTGATTCGGTTCAGAAACCAAGCAGAGTGGCAGCCGCAAGAGTCCCTTTTA 1683
QY 1501 TCTCCACTGATCGAAGTGGAGACGTGAGAGCAATTAACACACAGCTTCTTTTCGG 1560
Db 1684 TCTCCACTGATCGAAGTGGAGACGTGAGAGCAATTAACACACAGCTTCTTTTCGG 1743
QY 1561 CTTTCCACAGAGCAGAGACCTCAGAGAACTGTGCTGCGCTGGGCTTAAAGGCTGCGAC 1620
Db 1744 CTTTCCACAGAGCAGAGACCTCAGAGAACTGTGCTGCGCTGGGCTTAAAGGCTGCGAC 1803
QY 1621 TCGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGAGCTGTTTGGCC 1680
Db 1804 TCGGATATCTTGGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTTTGGCC 1863
QY 1681 ACAGAGTCTCACTTCTTACTCTGCTGAGCCATCTACGAGAGAGAGTCCAGTACTCT 1740
Db 1864 ACAGAGTCTCACTTCTTACTCTGCTGAGCCATCTACGAGAGAGAGTCCAGTACTCT 1923
QY 1741 GCCTACAGCTGACGACGCTGCCACTTGGCGAGACCAAGTCTATTTCTGTGCGCAGGCGG 1800

Db 1924 GCGTACAGCTGCGAGCTGCGCACTTGGGAGACCAAGTCTATTCTGTGCGCAGCGG 1983
Qy 1801 CAGAGCGAAGTACAGAGCTGACTGCGCGCGAGCTGGCATGAAGAAGCCCTTTGAA 1860
Db 1984 CAGAGCGAAGTACAGAGCTGACTGCGCGCGAGCTGGCATGAAGAAGCCCTTTGAA 2043
Qy 1861 AAGCAGTTTAAACGAGAGCTGCCAATGAAATTTGGAGAGCATCATGTCCAGAGAC 1920
Db 2044 AAGCAGTTTAAACGAGAGCTGCCAATGAAATTTGGAGAGCATCATGTCCAGAGAC 2103
Qy 1921 AGGTACCGGGAAGAGCTGGGGAAAGTGGGCACTGCTAGCTTTTCCGGCAGAGTGA 1980
Db 2104 AGGTACCGGGAAGAGCTGGGGAAAGTGGGCACTGCTAGCTTTTCCGGCAGAGTGA 2163
Qy 1981 ATCATTGAGGTCTCC 1995
Db 2164 ATCATTGAGGTCTCC 2178

RESULT 10

US-10-296-115-520
; Sequence 520, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT FILING DATE: US/10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 520
; LENGTH: 2966
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2966)
; OTHER INFORMATION: n = a,t,c or g
US-10-296-115-520

Query Match 92.7%; Score 1849; DB 13; Length 2966;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 1949; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 45 GGTGGCTCTGCTGGAAGTGGAAACGAAAGAGTGTCTAATTGATAGCCGCCATTTGT 104
Db 67 GGTGGCTCTGCTGGAAGTGGAAACGAAAGAGTGTCTAATTGATAGCCGCCATTTGT 126
Qy 105 GGAATACATATATCCCAATTTTGGAGCCATTATATCACTGCTCAAGCTTATGAA 164
Db 127 GGAATACATATATCCCAATTTTGGAGCCATTATATCACTGCTCAAGCTTATGAA 186
Qy 165 GCGAAGTTGCAACGAGCAAGAGTGTATTAACAAGCTATCCGCACTTACGGAACA 224
Db 187 GCGAAGTTGCAACGAGCAAGAGTGTATTAACAAGCTATCCGCACTTACGGAACA 246
Qy 225 TAAAGTTGACATTAATGCAAGTCAAGAGTGTATTAACAAGCTATCCGCAAGTGT 284
Db 247 TAAAGTTGACATTAATGCAAGTCAAGAGTGTATTAACAAGCTATCCGCAAGTGT 306
Qy 285 TGCCTCTCTCTTGAAGTGTATTTCTACTGTACTTGTGGTAACTGGAAGAGCTT 344
Db 307 TGCCTCTCTCTTGAAGTGTATTTCTACTGTACTTGTGGTAACTGGAAGAGCTT 366
Qy 345 CAACCTCTGTTCACTGTTGAGAGTGGTGTGAGAGTCTCTGCTGTTTCCCTGAGCT 404
Db 367 CAACCTCTGTTCACTGTTGAGAGTGGTGTGAGAGTCTCTGCTGTTTCCCTGAGCT 426

Qy 405 CTGTGAAGAAATTCACCTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACTGTTGC 464
Db 427 CTGTGAAGAAATTCACCTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACTGTTGC 486
Qy 465 CAACATTTGGGCCCAACCCGAATTTCTTCCAACTTTATCTTGGCTCCAGCAGATGTCT 524
Db 487 CAACATTTGGGCCCAACCCGAATTTCTTCCAACTTTATCTTGGCTCCAGCAGATGTCT 546
Qy 525 CAACAAGAGCTGATACAGCAGAAATGGATGTATATGTGTTAAATGCCAGTATACCTG 584
Db 547 CAACAAGAGCTGATACAGCAGAAATGGATGTATATGTGTTAAATGCCAGATACCTG 606
Qy 585 TCCAAAGCTGACTTATATCCCGAGTCTCATTTTCCGCTGCTGCTGGAATGACAGCTT 644
Db 607 TCCAAAGCTGACTTATATCCCGAGTCTCATTTTCCGCTGCTGCTGGAATGACAGCTT 666
Qy 645 TTGTGAGAAATTTTGGCCGTGTTGGAACAATTCAGTATTTCAATGAAAGCAAAAGC 704
Db 667 TTGTGAGAAATTTTGGCCGTGTTGGAACAATTCAGTATTTCAATGAAAGCAAAAGC 726
Qy 705 CTCGAATGATGTGTTCTAGTGCATGTTTGTAGTGGGATCTCCGCTCCGCAACATGCG 764
Db 727 CTCGAATGATGTGTTCTAGTGCATGTTTGTAGTGGGATCTCCGCTCCGCAACATGCG 786
Qy 765 TATGCCCTACATCATGAAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGA 824
Db 787 TATGCCCTACATCATGAAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGA 846
Qy 825 AGAAAAAGACCTCTATATCTCCAACTTCATTTTCTGGGCAACTCTGGAATATGA 884
Db 847 AGAAAAAGACCTCTATATCTCCAACTTCATTTTCTGGGCAACTCTGGAATATGA 906
Qy 885 GAAAGATTTAAGAACCAAGACTGAGAGTCAAGGCCAAGAGCAAACTCAAGCTGTGCA 944
Db 907 GAAAGATTTAAGAACCAAGACTGAGAGTCAAGGCCAAGAGCAAACTCAAGCTGTGCA 966
Qy 945 CCTGGAAGAGCAAAATGACCTGTCCCTGCTCTCAAGAGTGTGACAGAAAGCCAGAC 1004
Db 967 CCTGGAAGAGCAAAATGACCTGTCCCTGCTCTCAAGAGTGTGACAGAAAGCCAGAC 1026
Qy 1005 GCCCTCAGTCCACCTGTGCGCACTCTGTAACCTCAGAGGACGAGCAAAAGCCGCT 1064
Db 1027 GCCCTCAGTCCACCTGTGCGCACTCTGTAACCTCAGAGGACGAGCAAAAGCCGCT 1086
Qy 1065 GAATCCCGCAGAGGTGCGCAGAGTGTGCGCAGCTGACAGCCGTGCTTTAGAGAACGCC 1124
Db 1087 GAATCCCGCAGAGGTGCGCAGAGTGTGCGCAGCTGACAGCCGTGCTTTAGAGAACGCC 1146
Qy 1125 GCTGTACAGGCGCTCAGTGGGCTGCACTGTCCGAGACAGGCTGGAAGACAGCAATTA 1184
Db 1147 GCTGTACAGGCGCTCAGTGGGCTGCACTGTCCGAGACAGGCTGGAAGACAGCAATTA 1206
Qy 1185 GCTCAAGCGTTCTTCTCTGTGATATCAATCAGTTTCATTTAGCCAGCAGTGGCAGC 1244
Db 1207 GCTCAAGCGTTCTTCTCTGTGATATCAATCAGTTTCATTTAGCCAGCAGTGGCAGC 1266
Qy 1245 ATCTTAATATGAGCTTCTCTCATCAGAAAGTCTTGAATATCAAACTTTCCACTAC 1304
Db 1267 ATCTTAATATGAGCTTCTCTCATCAGAAAGTCTTGAATATCAAACTTTCCACTAC 1326
Qy 1305 TCTGTATGGGACCAACAGCTATGCAAGTCTCCCTGTTGAGAACTATGAGAGCAGAC 1364
Db 1327 TCTGTATGGGACCAACAGCTATGCAAGTCTCCCTGTTGAGAACTATGAGAGCAGAC 1386
Qy 1365 TCCGGAACCGAGTCTGTATTAAGAGAGAGCCAGCATCTCCCAAGAGTGCAGACGCGCAG 1424
Db 1387 TCCGGAACCGAGTCTGTATTAAGAGAGAGCCAGCATCTCCCAAGAGTGCAGACGCGCAG 1446
Qy 1425 GCGTTCAAGACGCGAGACAGATTCATTTGGGTGAGAAACGAGCAGTGGCACCGC 1484
Db 1447 GCGTTCAAGACGCGAGACAGATTCATTTGGGTGAGAAACGAGCAGTGGCACCGC 1506
Qy 1485 CCAAGAGTCCCTTTATCTCAGTGCATGCAAGTGGAGCGGTGAGAGCAATTAACAGAC 1544

Db	1507	CCAGAGGCTCCCTTTATATCTCCATCTGACATGGAAGTGGAGCGTGGAGCAATTATCCACAC	1566
Qy	1545	CAGCTTCCTTTTCGGCTCTTTCACACAGCCAGACAGCACTCAAGAACTCTGCTGGCTGGG	1604
Db	1567	CAGCTTCCTTTTCGGCTCTTTCACACAGCCAGACAGCACTCAAGAACTCTGCTGGCTGGG	1626
Qy	1605	CCTTAAGGGCTGGACCTCGGATATCTTGAGCCCCCAGACTCTTACCCCTTCCTGACCG	1664
Db	1627	CCTTAAGGGCTGGACCTCGGATATCTTGAGCCCCCAGACTCTTACCCCTTCCTGACCG	1686
Qy	1665	CAGCTGATATTTTGGCAGAGTCCCTACACTTCTACTCTGCTCAGCATCTACGGAGG	1724
Db	1687	CAGCTGATATTTTGGCAGAGTCCCTACACTTCTACTCTGCTCAGCATCTACGGAGG	1746
Qy	1725	CAGTGCAGTTACTCTCTGCTCAAGCTGCAGCCAGCTGCCACTTTCGGAGACCAAGTCTA	1784
Db	1747	CAGTGCAGTTACTCTCTGCTCAAGCTGCAGCCAGCTGCCACTTTCGGAGACCAAGTCTA	1806
Qy	1785	TTCTGTGGCGCAGGCGCGCAGAACCAAGTGAAGAAGCTACTCTGCGGCGGAGCTGCATGA	1844
Db	1807	TTCTGTGGCGCAGGCGCGCAGAACCAAGTGAAGAAGCTACTCTGCGGCGGAGCTGCATGA	1866
Qy	1845	AGAGAGCCCCCTTTGAAAAGCAGTTTAAACGACAGAACTGCCAAATGTGAATTTGGAGAG	1904
Db	1867	AGAGAGCCCCCTTTGAAAAGCAGTTTAAACGACAGAACTGCCAAATGTGAATTTGGAGAG	1926
Qy	1905	CATCATGTCAAGAAACAGGTCAACGGGAAGAGCTGGGAAAAGTGGCACTAGTCTAGCTT	1964
Db	1927	CATCATGTCAAGAAACAGGTCAACGGGAAGAGCTGGGAAAAGTGGCACTAGTCTAGCTT	1986
Qy	1965	TTTCGGGCGCATGGAATCATTTAGAGTCTCC	1995
Db	1987	TTTCGGGCGCATGGAATCATTTAGAGTCTCC	2017

```

RESULT 11
US-10-168-506-2
/ Sequence 2, Application US/10168506
/ Publication No. US20040053229a1
/ GENERAL INFORMATION:
/ APPLICANT: PLOWMAN, GREGORY D.
/ APPLICANT: MARTINEZ, RICARDO
/ APPLICANT: WHITE, DAVID
/ APPLICANT: MANNING, GERRAD
/ APPLICANT: SUPARSANAW, SUCHA
/ APPLICANT: HILL, RON
/ APPLICANT: FLANAGAN, PETER
/ TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
/ FILE REFERENCE: 038602/1351
/ CURRENT APPLICATION NUMBER: US/10/168,506
/ CURRENT FILING DATE: 2002-06-21
/ PRIOR APPLICATION NUMBER: PCT/US00/34736
/ PRIOR FILING DATE: 2000-12-21
/ NUMBER OF SEQ ID NOS: 76
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 2732
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-168-506-2

```

Query Match	92.3%	Score 1842;	DB 13;	Length 2732;
Best local Similarity	99.8%;	Pred. No. 0;		
Matches 1992; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0.

Qy 1 ATGGCCCATGATGTGTAATGCAACTCAATTTGTACTGAAAGCTGGCTCTGCTGCAA 60
Db 538 ATGGCCCATGAGTGTGTAATGCAATTTGTACTGAAAGCTGGCTCTGCTGCAA 597
Qy 61 AGTGAACGGAAATAAGTGTGCTAATTGATAGCGGCGCAATTTGTGSAATACATATACCC 120
Db 598 AATGGAACGGAAATAAGTGTGCTAATTGATAGCGGCGCAATTTGTGSAATACATATACCC 657

QY	121	CACATTTTGGAGCCATTAATATCACTGCTCCAACTTTATGAAGCGAAGTTGCACACG	180
Db	658	CACATTTTGGAGCCATTAATATCACTGCTCCAACTTTATGAAGCGAAGTTGCACACG	717
QY	181	GACAAAGTTAATTAACAGAGCTCATCAGCATTCAGGAAACATAAGTTGACATTGAT	240
Db	718	GACAAAGTTAATTAACAGAGCTCATCAGCATTCAGGAAACATAAGTTGACATTGAT	777
QY	241	TGCAGTCGAAGAGTTGTGTAGTTTACGATCAAAAGCTCCCAAGATGTGTGCTCTCTCTTCA	300
Db	778	TGCAGTCGAAGAGTTGTGTAGTTTACGATCAAAAGCTCCCAAGATGTGTGCTCTCTCTTCA	837
QY	301	GACTGTTTCTCACTGTATCTTGTGGGTAACTGAGAAAGAGCTTCAACTCTGTTACCTG	360
Db	838	GACTGTTTCTCACTGTACTTCTGGGTAACTGAGAAAGAGCTTCAACTCTGTTACCTG	897
QY	351	CTTGCAAGGTGGGTTTGTGAGATTCTCTCGTTGTTTCCCTGGCCCTGTGAAAGAAATCC	420
Db	898	CTTGCAAGGTGGGTTTGTGAGATTCTCTCGTTGTTTCCCTGGCCCTGTGAAAGAAATCC	957
QY	421	ACTGAGAGCCCTACCTGCATTTCTCAGAGCTTCACTGTTGCAACATTTGGAGCCAAAC	480
Db	958	ACTGAGAGCCCTACCTGCATTTCTCAGAGCTTCTTCACTGTTGCAACATTTGGAGCCAAAC	1017
QY	481	GCAATTTCTCCCAATCTTTATCTTGTGCTGCCAGAGAGATGTCCTCAACAAGAGCTGATA	540
Db	1018	GCAATTTCTCCCAATCTTTATCTTGTGCTGCCAGAGAGATGTCCTCAACAAGAGAGCTGATG	1077
QY	541	CAGCAGAAATGGGATTTGGTTATGTGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTT	600
Db	1078	CAGCAGAAATGGGATTTGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTT	1137
QY	601	ATCCCCAGAGTCATTTCTGCGTGGCTGTGAAATGACAGCTTTGTGGAAGAAATTTTG	660
Db	1138	ATCCCCAGAGTCATTTCTGCGTGGCTGTGAAATGACAGCTTTGTGGAAGAAATTTTG	1197
QY	661	CCGTGGTTGGCAAAATCAGTAGATTTCTATGAGAAGCAAAAGCCTCCAAATGATGTGT	720
Db	1198	CCGTGGTTGGCAAAATCAGTAGATTTCTATGAGAAGCAAAAGCCTCCAAATGATGTGT	1257
QY	721	CTAGTGCACTGTTAAGCTGGGATCTCCCGCTCCGCACCATCGCTATCGCTTACATCATG	780
Db	1258	CTAGTGCACTGTTAAGCTGGGATCTCCCGCTCCGCACCATCGCTATCGCTTACATCATG	1317
QY	781	AAAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGGAAGAAAGAAAGACCTACT	840
Db	1318	AAAGAGATGACATGTCTTTAGATGAAGCTTACAGATTTGGAAGAAAGAAAGACCTACT	1377
QY	841	ATATCTCCAAACTTCAATTTTCTGGGCAACTCTCGAGCTATGAGAAGATTAAAGAC	900
Db	1378	ATATCTCCAAACTTCAATTTTCTGGGCAACTCTCGAGCTATGAGAAGATTAAAGAC	1437
QY	901	CAGATGAGAGCATCAGGGCCAAAGAGCAAACTCAAGTGTCTGCACTTGGAGAGCCAAAT	960
Db	1438	CAGATGAGAGCATCAGGGCCAAAGAGCAAACTCAAGTGTCTGCACTTGGAGAGCCAAAT	1497
QY	961	GAACTGTGCTCTGCTGTCTCAGAGGTGTGACAAAAAGCGAGAGAGCCCTCAGTCCACC	1020
Db	1498	GAACTGTGCTCTGCTGTCTCAGAGGTGTGACAAAAAGCGAGAGAGCCCTCAGTCCACC	1557
QY	1021	TGTGCGCACTCTGCTACTCAGAGGACAGAGCAAAAGGCCCGTGCATATCCCGCAGAGGTG	1080
Db	1558	TGTGCGCACTCTGCTACTCAGAGGACAGAGCAAAAGGCCCGTGCATATCCCGCAGAGGTG	1617
QY	1081	CCAGAGCGTCCAGAGCTGACAGCCGTGTGTTAGAGGACAGCCCGCTGTGACAGGCGCTC	1140
Db	1618	CCAGAGCGTCCAGAGCTGACAGCCGTGTGTTAGAGGACAGCCCGCTGTGACAGGCGCTC	1677
QY	1141	AGTGGGCTGCACCTGTCCGACAGACAGGCTGTGAAGACACAGATAAGCTCAAGCTTCTTCC	1200
Db	1678	AGTGGGCTGCACCTGTCCGACAGACAGGCTGTGAAGACACAGATAAGCTCAAGCTTCTTCC	1737

QY 1201 TCTGTGATATCAATACGTTTATATATGACGAGATGGACATCTTTAATAGTGGCTTC 1260
DB 1738 TCTGTGATATCAATACGTTTATATATGACGAGATGGACATCTTTAATAGTGGCTTC 1797
QY 1261 TCTGTATCAGAGATGCTTTGGAATACTACAAACCTTCACATCTGTGATGGAGCAAC 1320
DB 1798 TCTGTATCAGAGATGCTTTGGAATACTACAAACCTTCACATCTGTGATGGAGCAAC 1857
QY 1321 AAGCTATGCGAGTTCCTCCCTGTGAGAACTATGAGAGAGATCTCCGAAACGATGCT 1380
DB 1858 AAGCTATGCGAGTTCCTCCCTGTGAGAACTATGAGAGAGATCTCCGAAACGATGCT 1917
QY 1381 GATAGAGAGAAAGCCAGATCTCCCAAGAGCTGACAGCCGAGGCTTTGACAGAGCAG 1440
DB 1918 GATAGAGAGAAAGCCAGATCTCCCAAGAGCTGACAGCCGAGGCTTTGACAGAGCAG 1977
QY 1441 AAGCAAGCATGTCATTCGGTCAAGAACGACAGCACTGCAACGAGCTTCCTTTTGA 1500
DB 1978 AAGCAAGCATGTCATTCGGTCAAGAACGACAGCACTGCAACGAGCTTCCTTTTGA 2037
QY 1501 TCTGCATGATCGAAGTGGAGCGTGGAGCAATTCACACAGCTTCCTTTTGGC 1560
DB 2038 TCTGCATGATCGAAGTGGAGCGTGGAGCAATTCACACAGCTTCCTTTTGGC 2097
QY 1561 CTTTCCACAGCCAGACAGACCTCAAGAGTCTGCTGGCTGGCTTTAAAGGCTGGCAC 1620
DB 2098 CTTTCCACAGCCAGACAGACCTCAAGAGTCTGCTGGCTGGCTTTAAAGGCTGGCAC 2157
QY 1621 TCGATATCTTGAGCCGCCAGACCTCAACCTTCCTCCGACAGCACTGCTTTTGGC 1680
DB 2158 TCGATATCTTGAGCCGCCAGACCTCAACCTTCCTCCGACAGCACTGCTTTTGGC 2217
QY 1681 ACAGAGTCTCAACCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
DB 2218 ACAGAGTCTCAACCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2277
QY 1741 GCTTACAGCTGCAAGCTGCGCCCACTTGGGAGAGCCAACTTATTTCTGCTGCTGCTG 1800
DB 2278 GCTTACAGCTGCAAGCTGCGCCCACTTGGGAGAGCCAACTTATTTCTGCTGCTGCTG 2337
QY 1801 CAGAGCGAAGTGAAGAGTGAAGTGGCGGAGCTGAGTGAAGAGAGCCCTTTGAA 1860
DB 2338 CAGAGCGAAGTGAAGAGTGAAGTGGCGGAGCTGAGTGAAGAGAGCCCTTTGAA 2397
QY 1861 AAGCAAGTTTAAAGCGAGAGCTGCCAAATGGAATTTGAGAGAGCATGTCAGAGAAC 1920
DB 2398 AAGCAAGTTTAAAGCGAGAGCTGCCAAATGGAATTTGAGAGAGCATGTCAGAGAAC 2457
QY 1921 AAGTCAAGGAGAGAGCTGGGGAAGTGGCAAGTCAAGCTTTTGGGGGAGCATGGA 1980
DB 2458 AAGTCAAGGAGAGAGCTGGGGAAGTGGCAAGTCAAGCTTTTGGGGGAGCATGGA 2517
QY 1981 ATGATGAGAGTCTCC 1995
DB 2518 ATGATGAGAGTCTCC 2532

RESULT 12

US-10-094-749-673
; Sequence 673, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI

APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHITO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
FILE REFERENCE: 08435/0160
CURRENT APPLICATION NUMBER: US/10/094,749
PRIOR FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO: 673
LENGTH: 2102
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-673

Query Match 89.8%; Score 1791; DB 16; Length 2102;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGCCCATAGAGATGATGGAACCTCAATGTTTATCTGAGAGTGGCTGCTGCTGGAA 60
DB 56 ATGGCCCATAGAGATGATGGAACCTCAATGTTTATCTGAGAGTGGCTGCTGCTGGAA 115
QY 61 AGTGAACGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 116 AGTGAACGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 175
QY 121 CACATTTTGGAGGATTAATATCACTGCTCCAGGTTATGAAGGAGTGGCAACG 180
DB 176 CACATTTTGGAGGATTAATATCACTGCTCCAGGTTATGAAGGAGTGGCAACG 235
QY 181 GACAAAGTTTAAATTAACAGAGCTCATCCAGCATTCAGGAAACATAGTTGACATTGAT 240
DB 236 GACAAAGTTTAAATTAACAGAGCTCATCCAGCATTCAGGAAACATAGTTGACATTGAT 295
QY 241 TGACGTGAGAGGTTGATGTTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 300
DB 296 TGACGTGAGAGGTTGATGTTTACATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 355
QY 301 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAAAGCTTCAACTCTGTCATG 360
DB 356 GACTGTTTCTCACTGATCTTCTGAGTAACTGAGAAAGCTTCAACTCTGTCATG 415
QY 361 CTTCAGAGGAGGTTGCTGATGTTCTGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 416 CTTCAGAGGAGGTTGCTGATGTTCTGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTG 475
QY 421 ACTTAGTCCCTACCTGATCTTCTGAGCTTCTTACCTGTTCCAAATTTGGGCAAC 480
DB 476 ACTTAGTCCCTACCTGATCTTCTGAGCTTCTTACCTGTTCCAAATTTGGGCAAC 535
QY 481 CGAATTTCTCCAACTTTATCTTGGCTGCGAGCGAGATGTCCTCAACAGAGACTGAT 540
DB 536 CGAATTTCTCCAACTTTATCTTGGCTGCGAGCGAGATGTCCTCAACAGAGACTGAT 595
QY 541 CAGCAAGATGGATGTTGTTATGTTAAATGCGAGCTTATCTGCTCCAAAGCTGACTTT 600
DB 596 CAGCAAGATGGATGTTGTTATGTTAAATGCGAGCTTATCTGCTCCAAAGCTGACTTT 655
QY 601 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 656 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 715
QY 661 CCGTGGTTGACAAATCAGTATGTTTATTTGAGAGCAAAAGCTTCAATGATGTTT 720

Db 716 CCGTGGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCATGAGTAGTGTT 775
Qy 721 CTAGTGCATCTTTTACTTGGGATCTCCCGTCCGCCACCATGCTATGCGCTACATCAG 780
Db 776 CTAGTGCATCTTTTACTTGGGATCTCCCGTCCGCCACCATGCTATGCGCTACATCAG 835
Qy 781 AAGAGATGGAACATGCTTTAGATGAACTTACAGATTTGTGAAGAAAAAGACCTACT 840
Db 836 AAGAGATGGAACATGCTTTAGATGAACTTACAGATTTGTGAAGAAAAAGACCTACT 895
Qy 841 ATATCTCAAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAAGAC 900
Db 896 ATATCTCAAACTTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAAAGAC 955
Qy 901 CAGACTGGAGCATCAGGGCCAAAGAGAACTCAGCTGCTGACCTGGAGAAAGCCAAAT 960
Db 956 CAGACTGGAGCATCAGGGCCAAAGAGAACTCAGCTGCTGACCTGGAGAAAGCCAAAT 1015
Qy 961 GAACTGTCCCTGCTGCTCAGAGGGTGAAGAAAAAGGAGAGGCGCCCTCAGTCCACCC 1020
Db 1016 GAACTGTCCCTGCTGCTCAGAGGGTGAAGAAAAAGGAGAGGCGCCCTCAGTCCACCC 1075
Qy 1021 TGTGCCGACTCTGCTACCTCAGAGGAGAGCAAAAGGCCGTGATCCCGCAGCGTG 1080
Db 1076 TGTGCCGACTCTGCTACCTCAGAGGAGAGCAAAAGGCCGTGATCCCGCAGCGTG 1135
Qy 1081 CCCAGCGTCCCGAGCGTGCAGCGCTGTGAGAGACAGCCCGCTGGTACAGGCGCTC 1140
Db 1136 CCCAGCGTCCCGAGCGTGCAGCGCTGTGAGAGACAGCCCGCTGGTACAGGCGCTC 1195
Qy 1141 AGTGGGCTGACCTGTCCGAGACAGGCTGGAGAGACAGCAATTAAGCTCAAGCGTTCCTC 1200
Db 1196 AGTGGGCTGACCTGTCCGAGACAGGCTGGAGAGACAGCAATTAAGCTCAAGCGTTCCTC 1255
Qy 1201 TCTCTGATATCAAAATCAGTTTCATATTCAGCCAGCATGCGACATCTTACATGCTTC 1260
Db 1256 TCTCTGATATCAAAATCAGTTTCATATTCAGCCAGCATGCGACATCTTACATGCTTC 1315
Qy 1261 TCTCTATCAGAGATGCTTTGGAATCTACAACTTCCACTATCTGTGATGGAGCAAC 1320
Db 1316 TCTCTATCAGAGATGCTTTGGAATCTACAACTTCCACTATCTGTGATGGAGCAAC 1375
Qy 1321 AAGCTATGCGAGTTCTCCCGTTCAGGAACATTCGAGAGAGCTCCGAAACCAAGTCT 1380
Db 1376 AAGCTATGCGAGTTCTCCCGTTCAGGAACATTCGAGAGAGCTCCGAAACCAAGTCT 1435
Qy 1381 GATTAAGAGAAAGCCAGCATCCCAAGAGCTGACAGCGCCAGGCTTCAGACAGCAG 1440
Db 1436 GATTAAGAGAAAGCCAGCATCCCAAGAGAGCTGACAGCGCCAGGCTTCAGACAGCAG 1495
Qy 1441 AGCAGCGATTGCTGGTCAAGACCAAGCAGAGTGGAACCGGCCAAGAGTCCCTTTTA 1500
Db 1496 AGCAGCGATTGCTGGTCAAGACCAAGCAGAGTGGAACCGGCCAAGAGTCCCTTTTA 1555
Qy 1501 TCTCCACTGATGAGAGTGGAGCGGTGGAGAAATTAACCAACAGCTTCCTTTCCGGC 1560
Db 1556 TCTCCACTGATGAGAGTGGAGCGGTGGAGAAATTAACCAACAGCTTCCTTTCCGGC 1615
Qy 1561 CTTTCAACAGCAGAGCACTCAGAAAGTGTGAGCTGGGCTTAAAGGGCTGGAC 1620
Db 1616 CTTTCAACAGCAGAGCACTCAGAAAGTGTGAGCTGGGCTTAAAGGGCTGGAC 1675
Qy 1621 TCGGATATCTTGGCCGCCAGACCTTACCTTCCCTGACAGAGCTGTATTTTGGC 1680
Db 1676 TCGGATATCTTGGCCGCCAGACCTTACCTTCCCTGACAGAGCTGTATTTTGGC 1735
Qy 1681 ACAGAGTCTCAGACTTCTACTGCTCAGCCATCTACGAGAGGAGTGCAGTTACTCT 1740
Db 1736 ACAGAGTCTCAGACTTCTACTGCTCAGCCATCTACGAGAGGAGTGCAGTTACTCT 1795
Qy 1741 GCTTACAGCTGACAGCAGCTGCCACTTGGAGAGCAAAAGTCTATCTGTGCGAGGCG 1800
Db 1796 GCTTACAGCTGACAGCAGCTGCCACTTGGAGAGCAAAAGTCTATCTGTGCGAGGCG 1855

Qy 1801 CAGAAAGCAAGTGAACAGAGCTGACTCGGCGGAGAGCTGGCATGAAGAGAGCCCTTTGA 1860
Db 1856 CAGAAAGCAAGTGAACAGAGCTGACTCGGCGGAGAGCTGGCATGAAGAGAGCCCTTTGA 1915
Qy 1861 AAGCATTTAAAGCAGAAAGCTGCCAAATGAAATTTGAGAGAGCATGTCAGAGAA 1920
Db 1916 AAGCATTTAAAGCAGAAAGCTGCCAAATGAAATTTGAGAGAGCATGTCAGAGAA 1975
Qy 1921 AGGTACGAGAAAGAGTGGGAAAGTGGGAGTCACTTACCTTTTGGGAGCAGATGAA 1980
Db 1976 AGGTACGAGAAAGAGTGGGAAAGTGGGAGTCACTTACCTTTTGGGAGCAGATGAA 2035
Qy 1981 ATCATTGAGTCTCC 1995
Db 2036 ATCATTGAGTCTCC 2050

Search completed: June 22, 2004, 16:36:33
Job time : 1178 secs

This Page Blank (uspto)

US-09-816-494-3

576.
TCCAGCT

```
/ GENERAL INFORMATION:  
 / APPLICANT: Meyers, Rachel A.  
 / TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY  
 / FILE REFERENCE: PHOSPHATASE MOLECULES AND USES THEREFOR  
 / CURRENT APPLICATION NUMBER: US/09/816,494  
 / PRIOR FILING DATE: 2001-03-23  
 / PRIORITY CLAIMING NUMBER: US 60/191,858  
 / NUMBER OF SEQ ID NOS: 10  
 / SOFTWARE: FastSeq for Windows Version 4.0  
 / SEQ ID NO 1  
 / LENGTH: 3544  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / NAME/KEY: CDS  
 / LOCATION: (589)...(2583)  
 / US-09-816-494-1
```

Query Match 94.9%; Score 1893; DB 4; Length 3544;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 2; Indels 0; Gaps 0

OY					1 ATGGCCCATGAGTGAATTGGAACTCAAAATTGTACTGAGAGGTTGGTGCTGCCTCGAGA	60
Dd					589 ATGGCCCATGAGTGAATTGGAACTCAAAATTGTACTGAGAGGTTGGTGCTGCCTCGAGA	648
OY					61 AGTGAACCGAAAAAAGTCGCTTAATTGATAGCCGCCATTGTGTGAATAACAATCATC	120
Dd					649 AGTGAACCGAAAAGTCGCTTAATTGATAGCCGCCATTGTGTGAATAACAATCATC	708
OY					121 CATATTGGAAACCATTATATATCAAATCTCCAAGCTTATGAACGAAGTTGCAACG	180
Dd					709 CATATTGGAAACCATTATATATCAAATCTCCAAGCTTATGAACGAAGTTGCAACG	768
OY					181 GAACAAGTTAATTAACAGAGCTCACGACATTCAGCAAGAACATPAAGTTGACAT	240
Dd					769 GAACAAGTTAATTAACAGAGCTCACGACATTCAGCAAGAACATPAAGTTGACAT	828
OY					241 TGACGCACAAGGTTGATTTACATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA	300
Dd					829 TGACGCACAAGGTTGATTTACATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA	888
OY					301 GACTGTTTTCTCACTGACTCTCTGGGTAACATGAGAGAGACTTCACTCGTTACCTG	360
Dd					889 GACTGTTTTCTCACTGACTCTCTGGGTAACATGAGAGAGACTTCACTCGTTACCTG	948
OY					361 CTTCAGAGTGGGTTTGCTGAAGTTCTCTGTTGTTCCCTGGCCTCTGTGAAGAAAATCC	420
Dd					949 CTTCAGAGTGGGTTTGCTGAAGTTCTCTGTTGTTCCCTGGCCTCTGTGAAGAAAATCC	1008
OY					421 ACTTAGAGCCCACTGATTTCTCAAGCTTGCTTACTGTTGCAACATTGGGCCCAACC	480
Dd					1009 ACTTAGAGCCCACTGATTTCTCAAGCTTGCTTACTGTTGCAACATTGGGCCCAACC	1068
OY					481 CGAATTTCTCCAACTTTATCTTGGSGTCGACAGAGATGTCNCAACAGAGAGCTGATA	540
Dd					1069 CGAATTTCTCCAACTTTATCTTGGSGTCGACAGAGATGTCNCAACAGAGAGCTGATA	1128
OY					541 CAGAGAAATGGGATTTGGTATGTGTTAATGCAGAGCTATACCTGTCCAAAGCTGACTTT	600
Dd					1129 CAGAGAAATGGGATTTGGTATGTGTTAATGCAGAGAAATPACTGTCCAAAGCTGACTTT	1188
OY					601 ATCCCAGACTCAATTTCTGCGCGTGCCGTGAATGACAGCTTTGTGAGAAAATTTTG	660

Query Match	Best Local Similarity	94.9%;	Score 1893;	DB 4;	Length 3544;			
Matches 1993;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0
QY	1	ATGGCCCATGAGATGATTTGGAACTCAAAATTGTTACTGAGAGATGGTGGCTCTGCTGGAA	60					
Db	589	ATGGCCCATGAGATGATTTGGAACTCAAAATTGTTACTGAGAGATGGTGGCTCTGCTGGAA	648					
QY	61	AGTGAACGGAAAAGGTGCTGCTAAATTGGATTAGCCGGCCATTGTGGATATACATTC	120					
Db	649	AGTGAACGGAAAAGGTGCTGCTAAATTGGATTAGCCGGCCATTGTGGATATACATTC	708					
QY	121	CACATTTTGGAAACCCATTAATATCAACTCTCCAAAGCTTATGAACCGAAGTTGGACAG	180					
Db	709	CACATTTTGGAAACCCATTAATATCAACTCTCCAAAGCTTATGAACCGAAGTTGGACAG	768					
QY	181	GACAAAGTGTATTATACAGAGCTATCCAGATTCAGGAAACATAAGTTGACATTGAAT	240					
Db	769	GACAAAGTGTATTATACAGAGCTATCCAGATTCAGGAAACATAAGTTGACATTGAAT	828					
QY	241	TGCAAGTCAAGAGTTGATTTAGCATCAAAAGCTCCAAAGATGTTGCTCTCTCTTTCA	300					
Db	829	TGCAAGTCAAGAGTTGATTTAGCATCAAAAGCTCCAAAGATGTTGCTCTCTCTTTCA	888					
QY	301	GACTGTTTTCTACAGTACTCTGGGTAACTGAGAAAGAGCTTCAACTGTGTACCTG	360					
Db	889	GACTGTTTTCTACAGTACTCTGGGTAACTGAGAAAGAGCTTCAACTGTGTACCTG	948					
QY	361	CTTGCAGGTGGGTGTGCTGAGTCTCTCGTGTGTTTCCCTGGCCCTGTGAAGGAAATTC	420					
Db	949	CTTGCAGGTGGGTGTGCTGAGTCTCTCGTGTGTTTCCCTGGCCCTGTGTGAAGGAAATTC	1008					
QY	421	ACTTAAGTCCCTACCTGCATTTCTCAGCCCTTGCTTACCTGTGGCAACATTGGGCCAAC	480					
Db	1009	ACTTAAGTCCCTACCTGCATTTCTCAGCCCTTGCTTACCTGTGGCAACATTGGGCCAAC	1068					
QY	481	CGAATTTCTCCCAATCTTTATCTTGGGTGGCAGCGAGATGTCTCTCAACAGAGCTGATA	540					
Db	1069	CGAATTTCTCCCAATCTTTATCTTGGGTGGCAGCGAGATGTCTCTCAACAGAGCTGATG	1128					
QY	541	CAGCAGATGGGATTTGTTATGTGTTAAATGCGAGCATATCCGTGCCAAAGCCTGACTTT	600					
Db	1129	CAGCAGATGGGATTTGTTATGTGTTAAATGCGAGCATATCCGTGCCAAAGCCTGACTTT	1188					
QY	601	ATCCCCAGTCTCATTTTCTGCGGTGGCCTGTGAATGACAGCTTTTGTGAGAAAATTTTG	660					
Db	1189	ATCCCCAGTCTCATTTTCTGCGGTGGCCTGTGAATGACAGCTTTTGTGAGAAAATTTTG	1248					
QY	661	CCGTGTTTGAACAATCAGTAGTTTCATTGAGAAAGCAAAAGCCTCCATGATGTGTT	720					
Db	1249	CCGTGTTTGAACAATCAGTAGTTTCATTGAGAAAGCAAAAGCCTCCATGATGTGTT	1308					
QY	721	CTATGTGACTGTTTATGCTGGGATCTCCGCTCCGCCACCATATGAGCTATATGCTATATG	780					

Db 1309 CTAGTGCATGTTAGCTGGGATCTCCGCTCCGCCACATCGCTATCGCTCATCATG 1368
 Qy 781 AAGAGGATGGAACATGTTTATGATGAAGCTTACAGATTTGTGAAGAAAAAAGACTACT 840
 Db 1369 AAGAGGATGGAACATGTTTATGATGAAGCTTACAGATTTGTGAAGAAAAAAGACTACT 1428
 Qy 841 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTCTGACATATGAGAAAGATTTAAGAC 900
 Db 1429 ATATCTCCAAACTTCAATTTTCTGGGCAACTCTCTGACATATGAGAAAGATTTAAGAC 1488
 Qy 901 CAGACTGAGACATCAGGGCCAAAGAGCAACTCAAGCTGCTGCACTGGAAGCCAAAT 960
 Db 1489 CAGACTGAGACATCAGGGCCAAAGAGCAACTCAAGCTGCTGCACTGGAAGCCAAAT 1548
 Qy 961 GAACCTGTCCTCTGCTCTCAAGAGGTGGAACAGAAAAAGCAAGCCCTCAGTCCACC 1020
 Db 1549 GAACCTGTCCTCTGCTCTCAAGAGGTGGAACAGAAAAAGCAAGCCCTCAGTCCACC 1608
 Qy 1021 TGTGCGGACTCTGCTGCTCTCAAGAGGTGGAACAGAAAAAGCAAGCCCTCAGTCCACC 1080
 Db 1609 TGTGCGGACTCTGCTGCTCTCAAGAGGTGGAACAGAAAAAGCAAGCCCTCAGTCCACC 1668
 Qy 1081 CCCAGCTGCTCAGCTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 Db 1669 CCCAGCTGCTCAGCTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1728
 Qy 1141 AGTGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 Db 1729 AGTGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1788
 Qy 1201 TCTCTGATATCAATATGATTTATATATATATATATATATATATATATATATATATATAT 1260
 Db 1789 TCTCTGATATCAATATGATTTATATATATATATATATATATATATATATATATATATAT 1848
 Qy 1261 TCTCTATCAGAAAGTCTTGAAT 1320
 Db 1849 TCTCTATCAGAAAGTCTTGAAT 1908
 Qy 1321 AAGCTATGCAAGTTCTCCCTGTTTCAAGAACTATCGGAGCAACTCCGAAACCACTGCT 1380
 Db 1909 AAGCTATGCAAGTTCTCCCTGTTTCAAGAACTATCGGAGCAACTCCGAAACCACTGCT 1968
 Qy 1381 GATTAAGAGGAAGCTGCAATCTCCCAAGAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTG 1440
 Db 1969 GATTAAGAGGAAGCTGCAATCTCCCAAGAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTG 2028
 Qy 1441 AGCAAGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 Db 2029 AGCAAGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2088
 Qy 1501 TCTCCACTGATGGAAGTGGGAGGTGGAGGACAAATTACCAACCAAGCTTCTTTTGGC 1560
 Db 2089 TCTCCACTGATGGAAGTGGGAGGTGGAGGACAAATTACCAACCAAGCTTCTTTTGGC 2148
 Qy 1561 CTTTCCACGAGCAGCACTTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
 Db 2149 CTTTCCACGAGCAGCACTTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2208
 Qy 1621 TCGGATTTCTTGGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTTATTTTGGC 1680
 Db 2209 TCGGATTTCTTGGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGTTATTTTGGC 2268
 Qy 1681 ACAGAGTCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 Db 2269 ACAGAGTCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2328
 Qy 1741 GCTTACAGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 Db 2329 GCTTACAGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2388
 Qy 1801 CAGAAAGCAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860

Db 2389 CAGAAAGCAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2448
 Qy 1861 AAGCACTTTAAACGCAAGAGCTGCAATGGAATTTGAGAGAGCATCATGTGAGAGAAC 1920
 Db 2449 AAGCACTTTAAACGCAAGAGCTGCAATGGAATTTGAGAGAGCATCATGTGAGAGAAC 2508
 Qy 1921 AGTCAAGGGAAGAGCTGAGGGAAGTGGCAGTCACTTATGCTTTTGGGCAAGCATGGA 1980
 Db 2509 AGTCAAGGGAAGAGCTGAGGGAAGTGGCAGTCACTTATGCTTTTGGGCAAGCATGGA 2568
 Qy 1981 ATCATTGAGTCTCC 1995
 Db 2569 ATCATTGAGTCTCC 2583

Search completed: June 22, 2004, 16:16:35
 Job time: 173 secs

1995
 -575
 1340



This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 21, 2004, 13:25:03 ; Search time 21 Seconds
(without alignments)
3046.065 Million cell updates/sec

Title: US-10-029-345a-109

Perfect score: 665

Sequence: 1 MAHEMIGTQIVTERLVALL.....LGKVSQSFSGSMELIEVS 665

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 57676

Minimum DB seq length: 473

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database : PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.4	619	2	T15969
2	8	1.2	727	2	AB2353
3	8	1.2	1308	2	T15280
4	7	1.1	475	2	B70143
5	7	1.1	478	2	S17836
6	7	1.1	479	2	T47415
7	7	1.1	488	2	AE0001
8	7	1.1	507	2	G84577
9	7	1.1	509	2	D86911
10	7	1.1	509	2	T10013
11	7	1.1	510	2	A38337
12	7	1.1	511	2	AE3405
13	7	1.1	515	2	S61889
14	7	1.1	524	2	B97864
15	7	1.1	535	2	S40461
16	7	1.1	535	2	S40462
17	7	1.1	546	2	C72453
18	7	1.1	548	2	G81959
19	7	1.1	564	2	T72268
20	7	1.1	566	2	S50979
21	7	1.1	579	2	RT0494
22	7	1.1	583	2	T11624
23	7	1.1	585	2	S55205
24	7	1.1	596	2	T17333
25	7	1.1	601	2	T13054
26	7	1.1	615	2	G90575
27	7	1.1	622	2	B86431
28	7	1.1	628	1	A39262
29	7	1.1	628	1	A33333

30	7	1.1	634	2	G90779	hypothetical prote
31	7	1.1	657	2	S32991	hypothetical prote
32	7	1.1	657	2	F81896	probable phage tra
33	7	1.1	661	2	A81125	bacteriophage tran
34	7	1.1	661	2	S45131	probable membrane
35	7	1.1	674	2	C81505	hypothetical prote
36	7	1.1	680	2	E72033	hypothetical prote
37	7	1.1	680	2	B86590	hypothetical prote
38	7	1.1	693	2	G82618	pilus biogenesis p
39	7	1.1	700	1	S01991	transforming prote
40	7	1.1	704	1	S33704	transforming prote
41	7	1.1	731	2	T38117	probable protein k
42	7	1.1	704	2	T17002	probable beta-gala
43	7	1.1	745	2	C99995	hypothetical prote
44	7	1.1	745	2	H85840	hypothetical prote
45	7	1.1	757	2	E82013	probable transcrip
46	7	1.1	757	2	A81242	transcription acce
47	7	1.1	778	2	E97224	ATP-dependent lon
48	7	1.1	784	2	T51759	glutamate-ammonia
49	7	1.1	795	2	T34468	hypothetical prote
50	7	1.1	828	2	T03544	hypothetical prote
51	7	1.1	858	2	UC7683	lactate receptor TIR
52	7	1.1	866	2	F88481	protein C16A3.1 [1
53	7	1.1	896	2	D96556	hypothetical prote
54	7	1.1	930	2	T00403	hypothetical prote
55	7	1.1	936	2	S20480	crac protein - Esc
56	7	1.1	946	2	T38100	rho-GTPase-activat
57	7	1.1	966	2	G70838	probable tmpl1 pr
58	7	1.1	1052	2	T04439	hypothetical prote
59	7	1.1	1071	2	S48378	probable membrane
60	7	1.1	1099	2	A55405	adenylate cyclase
61	7	1.1	1115	2	T29012	hypothetical prote
62	7	1.1	1136	2	F96564	hypothetical prote
63	7	1.1	1165	2	D59433	C. elegans protein
64	7	1.1	1234	2	T00363	hypothetical prote
65	7	1.1	1264	2	B84534	probable retroelem
66	7	1.1	1277	2	T14152	synaptic scaffold
67	7	1.1	1309	2	S62457	hypothetical prote
68	7	1.1	1318	2	T38568	hypothetical prote
69	7	1.1	1454	2	E84535	probable retroelem
70	7	1.1	1461	2	E84589	probable retroelem
71	7	1.1	1465	2	T30891	PHY3 protein - mai
72	7	1.1	1642	2	I51018	cobra venom factor
73	7	1.1	1737	2	T00209	MEG8 protein - hu
74	7	1.1	1792	2	T08678	supervillin P205 -
75	7	1.1	2077	1	WZBZ24	240k tegument prot

ALIGNMENTS

RESULT 1
T15969
hypothetical protein F08B1.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C/Accession: T15969
R/Chisoe: S.
submitted to the EMBL Data Library, July 1995
A/Description: The sequence of C. elegans cosmid F08B1.
A/Reference number: Z18439
A/Accession: T15969
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-619 <CHI>
A/Cross-references: EMBL:U23178; NID:G726421; PID:G726422; PIDN:AMC46719.1; CESP:F08B1.1
A/Experimental source: strain Bristol N2
C/Genetic: S.
A/Gen: CESP:F08B1.1
A/Introns: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3
Query Match 1.4%; Score 9; DB 2; Length 619;
Best local Similarity 100.0%; Pred. No. 1.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 HCLAGISRS 251
|||||
Db 223 HCLAGISRS 231

RESULT 2
AB2353
hypothetical protein al14378 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AB2353
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kurlitz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Ref. 8, 205-213, 2001
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2353
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1727 <KIR>
A:Cross-references: GB:BA000019; PIDN:BA076077.1; PID:gl7133514; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: al14378

Query Match
Best Local Similarity 1.2%; Score 8; DB 2; Length 727;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 SVSPVPSV 366
|||||
Db 116 SVSPVPSV 123

RESULT 3
T15280
hypothetical protein R155.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15280
R:Giesel, C.; Wamsley, P.; Kramex, J.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid R155.
A:Reference number: Z18321
A:Accession: T15280
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1308 <GRI>
A:Cross-references: EMBL:AF003390; NID:g2088866; PID:g2088869; PIDN:AA054272.1; GSPDB:GN00179
A:Experimental source: strain Bristol N2; clone R155
C:Genetics:
A:Gene: CBSP.R155.2
A:Map position: 3
A:introns: 13/1; 57/1; 860/2; 897/1; 970/2; 1116/2; 1174/1; 1225/3

Query Match
Best Local Similarity 1.2%; Score 8; DB 2; Length 1308;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 247 GISRSATI 254
|||||
Db 1165 GISRSATI 1172

RESULT 4
B70143
fibronectin/fibrinogen-binding protein homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: B70143

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White son, D.; Peterson, J.; Keriavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A>Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: B70143
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-475 <KLE>
A:Cross-references: GB:AE001141; GB:AE000783; NID:g2688250; PIDN:AC66734.1; PID:g268825
A:Experimental source: strain B31

Query Match
Best Local Similarity 1.1%; Score 7; DB 2; Length 475;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 KKIKNOT 302
|||||
Db 469 KKIKNOT 475

RESULT 5
S17836
fatty-acyl-CoA reductase (EC 1.2.1.-) luxC - Photobacterium leiognathi
C:Species: Photobacterium leiognathi
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 26-May-2000
C:Accession: S17836
R:Lee, C.Y.; Sztitner, R.B.; Meighen, E.A.
Eur. J. Biochem. 201, 161-167, 1991
A>Title: The lux genes of the luminous bacterial symbiont, Photobacterium leiognathi, of coli.
A:Reference number: S17836; MUID:92007870; PMID:1915359
A:Accession: S17836
A:Molecule type: DNA
A:Residues: 1-478 <LEE>
A:Cross-references: EMBL:M63594; NID:g150687; PIDN:AAA25616.1; PID:g150688
C:Genetics:
A:Gene: luxC
C:Superfamily: Photobacterium fatty-acyl-CoA reductase
C:Keywords: oxidoreductase

Query Match
Best Local Similarity 1.1%; Score 7; DB 2; Length 478;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 SVHLIAG 123
|||||
Db 145 SVHLIAG 151

RESULT 6
T47415
transporter-like protein - Arabidopsis thaliana
N:Alternate names: protein T28A8 80
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47415
R:Punelle, B.; Boutry, M.; Goffeau, A.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24466
A:Accession: T47415
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <PUR>
A:Cross-references: EMBL:AL162691
A:Experimental source: cultivar Columbia; BAC clone T28A8
C:Genetics:
A:Map position: 3
A:introns: 48/1; 77/1; 111/2; 145/3; 163/3; 184/3; 201/2; 226/3; 291/3; 321/1; 349/3; 38
A:Note: T28A8.80


```
Query Match      1.1%; Score 7; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      518 LFGGLSTS 524
      |||||
Db      117 LFGGLSTS 123

RESULT 7
AE0001
conserved hypothetical protein YP00004 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #ext_change 09-Nov-2001
C/Accession: AE0001
R/Parhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
demo-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Accession: AE0001
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-488 <KUR>
A/Cross-references: GB:AL50842; PIDN:CAC88871.1; PID:g15978119; GSPDB:GN00175
C/Genetics:
A/Genes: YP00004
C/Suprafamily: Escherichia coli hypothetical 49.6K protein (asna 3' region)

Query Match      1.1%; Score 7; DB 2; Length 488;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      369 SLEEDSP 375
      |||||
Db      80 SLEEDSP 86

RESULT 8
G84577
probable WD-40 repeat protein, MS14 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #ext_change 02-Feb-2001
C/Accession: G84577
R/Jin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Xoo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: AB4420; MUID:20083487; PMID:10617197
A/Accession: G84577
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-507 <STO>
A/Cross-references: GB:AE002093; NID:g4191782; PIDN:AAD10151.1; GSPDB:GN00139
C/Genetics:
A/Genes: AC2919520
A/Map position: 2

Query Match      1.1%; Score 7; DB 2; Length 507;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      302 TGASGPK 308
      |||||
Db      20 TGASGPK 26

RESULT 9
D86911
conserved hypothetical protein ML0020 [imported] - Mycobacterium leprae
C/Species: Mycobacterium leprae
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #ext_change 20-Apr-2001
C/Accession: D86911
R/Cole, S.T.; Eigemeier, K.; Parhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H.
R.; Davies, R.M.; Devlin, K.; Duthey, S.; Felwell, T.; Fraser, A.; Hamlin, N.; Holroyd.
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squires, R.; Sk
A>Title: Massive gene decay in the leprosy bacillus.
A/Reference number: AB6909; MUID:21128732; PMID:11234002
A/Accession: D86911
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-509 <STO>
A/Cross-references: GB:AL450380; NID:g13092430; PIDN:CAC29528.1; GSPDB:GN00147
C/Genetics:
A/Genes: ML0020

Query Match      1.1%; Score 7; DB 2; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      489 TSSSGTA 495
      |||||
Db      460 TSSSGTA 466

RESULT 10
T10013
probable phosphoprotein phosphatase (EC 3.1.3.16) - Mycobacterium leprae
C/Species: Mycobacterium leprae
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #ext_change 16-Jul-1999
C/Accession: T10013
R/Cole, S.T.
Submitted to the EMBL Data Library, August 1997
A/Reference number: Z16916
A/Accession: T10013
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-509 <COL>
A/Cross-references: EMBL:Z70722; NID:e1059634; PID:e337960
C/Genetics:
A/Note: PPP
C/Keywords: phosphoric monoester hydrolase

Query Match      1.1%; Score 7; DB 2; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      489 TSSSGTA 495
      |||||
Db      460 TSSSGTA 466

RESULT 11
A38337
amido-phosphoribosyltransferase (EC 2.4.2.14) precursor - chicken
N/Alternate names: glutamine PRP amido-transferase
C/Species: Gallus gallus (chicken)
C/Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #ext_change 18-Feb-2000
C/Accession: A38337; I50240
R/Zhou, G.; Dixon, J.E.; Zalkin, H.
J. Biol. Chem. 265, 21152-21159, 1990
A>Title: Cloning and expression of avian glutamine phosphoribosylpyrophosphate amidotrai
A/Reference number: A38337; MUID:91065928; PMID:2123487
A/Accession: A38337
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-510 <ZHO>
A/Cross-references: GB:M60069; GB:M38186; NID:g211824; PIDN:AAA62736.1; PID:g211825
R/Gavala, A.; Dixon, J.E.; Braydon, K.A.; Zalkin, H.
Mol. Cell. Biol. 13, 4784-4792, 1993
A>Title: Coexpression of two closely linked avian genes for purine nucleotide synthesis
A/Reference number: I50239; MUID:93330273; PMID:8336716
```

A:Accession: I50240
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-42 <GAV>
A:Cross-references: GB:LI2533; NID:G402698; PIDN:AAA17895.1; PID:G482248
C:Genetics:
A:Gene: GPAT
C:Superfamily: amidophosphoribosyltransferase
C:Keywords: 4fe-4S; glycosyltransferase; metalloprotein; pentosyltransferase
F:1-11/Domain: propeptide #status predicted <PRO>
F:12-510/Product: amidophosphoribosyltransferase #status predicted <MAT>
F:12/Active site: Cys #status predicted
F:280,426,496,499/Binding site: 4fe-4S cluster (Cys) (covalent) #status predicted

Query Match 1.1%; Score 7; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 594 VVSRRR 600
|||||
DB 299 VVSRRR 305

RESULT 12
AE3405
hypothetical protein BME11227 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AE3405
R:Belicchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Lob, T.; Ivanova, P.; Mazur, M.; Goldsman, E.; Selkov, E.; Elker, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD252; PMID:11756688
A:Accession: AE3405
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-511 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52408.1; PID:G17983210; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11227
A:Map position: 1

Query Match 1.1%; Score 7; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 AKKVDI 79
|||||
DB 34 AKKVDI 40

RESULT 13
S61889
3-oxosteroid 1-dehydrogenase (EC 1.3.99.4) - Arthrobacter simplex
N:Alternate names: 3-ketosteroid-delta-dehydrogenase kedd
C:Species: Arthrobacter simplex
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Jun-2000
C:Accession: S61889
R:Mohar, I.; Choi, K.P.; Yamashita, M.; Murooka, Y.
Mol. Microbiol. 15, 895-905, 1995
A:Title: Molecular cloning, expression in Streptomyces lividans, and analysis of a gene se and a hypothetical regulatory protein.
A:Reference number: S61887; MUID:9511931; PMID:7596291
A:Accession: S61889
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-515 <MOU>
A:Cross-references: EMBL:D37969; NID:G882006; PIDN:BA07186.1; PID:G1518167
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
C:Genetics:

A:Gene: kedd
C:Superfamily: 3-oxosteroid 1-dehydrogenase; 3-oxosteroid 1-dehydrogenase homology
C:Keywords: oxidoreductase
F:403-501/Domain: 3-oxosteroid 1-dehydrogenase homology <OXD>

Query Match 1.1%; Score 7; DB 2; Length 515;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 EDALRY 430
|||||
DB 71 EDALRY 77

RESULT 14
B97864
hypothetical protein hemk [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: B97864
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: B97864
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-524 <KUR>
A:Cross-references: GB:AB06914; PIDN:AAL03852.1; PID:G15620455; GSPDB:GN00173
C:Genetics:
A:Gene: hemk

Query Match 1.1%; Score 7; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 ELIOQNG 184
|||||
DB 473 ELIOQNG 479

RESULT 15
S40461
t-complex-type molecular chaperone tcpl (clone ASTCP-K19) - oat
C:Species: Avena sativa (oat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999
C:Accession: S40461; S39319
R:Bhmann, B.; Krenz, M.; Mummert, E.; Schaefer, E.
FEBS Lett. 336, 313-316, 1993
A:Title: Two Tcpl-1-related but highly divergent gene families exist in oat encoding prot
A:Reference number: S40461; MUID:94085629; PMID:7903257
A:Accession: S40461
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-535 <EHM>
A:Cross-references: EMBL:X75777; NID:G435172; PIDN:CA53396.1; PID:G435173
A:Experimental source: cultivar pawi, tissue_type mesocotyl and coleoptile
C:Genetics:
A:Gene: tcpl
C:Superfamily: molecular chaperone t-complex-type
C:Keywords: molecular chaperone

Query Match 1.1%; Score 7; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 KHKVDI 80
|||||
DB 259 KHKVDI 265

RESULT 16
S40462

t-complex-type molecular chaperone tcp1 (clone ASTCP-K36) - oat
C:Species: *Avena sativa* (oat)
C>Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change.20-Aug-1999
C:Accession: S40462; S39320
R:Bhmann, B.; Krenz, M.; Mummert, E.; Schaefer, E.
FEBS Lett. 336, 313-316, 1993
A:Title: Two Tcp-1-related but highly divergent gene families exist in oat encoding prot
A:Reference number: S40461; MUID:94085629; PMID:7903257
A:Accession: S40462
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-535 <EMBL>
A:Cross-references: EMBL:X75778; NID:9435174; PIDN:CAA53397.1; PID:9435175
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C:Genetics:
A:Gene: tcp1
C:Superfamily: molecular chaperone t-complex-type
C:Keywords: molecular chaperone

Query Match 1.1% Score 7; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 KHKVDID 80
DB 259 KHKVDID 265

RESULT 17
C72453
hypothetical protein APE2271 - Aeropyrum pernix (strain K1)
C:Species: *Aeropyrum pernix*
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C:Accession: C72453
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyr*
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: C72453
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-546 <RAW>
A:Cross-references: DDBJ:AP000064; NID:95105945; PIDN:BAAB1283.1; PID:dl045069; PID:9510
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2271
C:Superfamily: Aeropyrum pernix hypothetical protein APE2271

Query Match 1.1% Score 7; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 567 YSASAIY 573
DB 373 YSASAIY 379

RESULT 18
G81959
conserved hypothetical protein NMA0428 [imported] - *Neisseria meningitidis* (strain Z2491
C:Species: *Neisseria meningitidis*
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: G81959
R:Fairhill, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churche, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: AB1775; MUID:20222556; PMID:10761919
A:Accession: G81959
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-548 <PAR>

A:Cross-references: GB:AL162753; GB:AL157959; NID:97379120; PIDN:CAB83727.1; PID:973791
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0428

Query Match 1.1% Score 7; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 562 ESSHYS 568
DB 281 ESSHYS 287

RESULT 19
F72268
ABC transporter, ATP-binding protein - *Thermotoga maritima* (strain MSB8)
C:Species: *Thermotoga maritima*
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: F72268
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: F72268
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-564 <ARN>
A:Cross-references: GB:AE001786; GB:AE000512; NID:94981873; PIDN:AAD36392.1; PID:949818
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TMW1319
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 1.1% Score 7; DB 2; Length 564;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 ESKSTLV 143
DB 367 ESKSTLV 373

RESULT 20
S50979
RHC21 protein - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein D2940; protein YD8119_04; protein YDL003w
C:Species: *Saccharomyces cerevisiae*
C>Date: 11-Feb-1995 #sequence_revision 12-May-1995 #text_change 05-Dec-1997
C:Accession: S50979; S55719; S52516; S67535
R:Murphy, L.; Richards, C.; Gentile, S.; Harris, D.
submitted to the EMBL Data Library, January 1995
A:Reference number: S50976
A:Accession: S50976
A:Molecule type: DNA
A:Residues: 1-566 <MUR>
A:Cross-references: EMBL:Z48008; NID:9642799; PID:9642803
R:Sturmiklov, A.V.; Koshland, D.
submitted to the EMBL Data Library, March 1995
A:Reference number: S55719
A:Accession: S55719
A:Molecule type: DNA
A:Residues: 1-566 <STR>
A:Cross-references: EMBL:U23759; NID:9777397; PID:9777398
R:Andre, B.; Vissers, S.; Urestarazu, L.
submitted to the EMBL Data Library, February 1995
A:Description: The sequence of a 42 kb segment located on the left arm of chromosome IV
A:Reference number: S52492
A:Accession: S52516
A:Molecule type: DNA
A:Residues: 1-285 <AND>

A;Cross-references: EMBL:Z48432; NID:9683669; PID:9683694
 A;Experimental source: strain S288C
 R;Ureastarazu, L.A.; Andre, B.; Vissers, S.
 submitted to the Protein Sequence Database, July 1996
 A;Reference number: 567535
 A;Accession: 567535
 A;Molecule type: DNA
 A;Residues: 1-285 <URR>
 A;Cross-references: EMBL:Z74051; MIPS:YDL003w
 A;Experimental source: strain S288C
 C;Genetics:
 A;Gene: SGD:MCD1; RHC21
 A;Cross-references: MIPS:YDL003w; SGD:S0002161
 A;Map position: 4L

Query Match 1.1%; Score 7; DB 2; Length 566;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 633 GESISMS 639
 Db 271 GESISMS 277

RESULT 21

JT0494
 alpha-glucosidase homolog precursor - yellow fever mosquito
 N;Alternate names: maltase homolog
 C;Species: Aedes aegypti (yellow fever mosquito)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 22-Jun-1999
 C;Accession: JT0494
 R;James, A.A.; Blackmer, K.; Racioppi, J.V.
 Gene 75, 73-83, 1989
 A;Title: A salivary gland-specific, maltase-like gene of the vector mosquito, Aedes aegypti
 A;Reference number: JT0494; MID:89252923; PMID:2470653
 A;Accession: JT0494
 A;Molecule type: mRNA
 A;Residues: 1-579 <UAM>
 A;Cross-references: GB:M23442; GB:M22322; GB:M22364; NID:9159565; PIDN:AAA29352.1; PID:9
 A;Note: the authors translated the codon CAA for residue 569 as Lys
 C;Genetics:
 A;Gene: Mali
 C;Superfamily: alpha-glucosidase; alpha-amylase core homology
 C;Keywords: glycoprotein
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-579/Product: alpha-glucosidase homolog #status predicted <MAT>
 F;187-359/Domain: alpha-amylase core homology <AMY>
 F;118,151,282,304,325,401/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 1.1%; Score 7; DB 2; Length 579;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 577 ASYSAYS 583
 Db 414 ASYSAYS 420

RESULT 22

T11624
 spindle poison sensitivity protein - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
 C;Accession: T11624
 R;Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, September 1997
 A;Reference number: Z17303
 A;Accession: T11624
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-583 <BAR>
 A;Cross-references: EMBL:Z99260; NID:e106197; PID:e1341400
 C;Genetics:

A;Map position: IL
 A;Introns: 35/1
 A;Note: SPAC3A11.02

Query Match 1.1%; Score 7; DB 2; Length 583;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 408 SYSASMA 414
 Db 298 SYSASMA 304

RESULT 23

S55205
 dihydroxy-acid dehydratase (EC 4.2.1.9) - yeast (Saccharomyces cerevisiae)
 N;Alternate names: protein J1450; protein YJR016c
 C;Species: Saccharomyces cerevisiae
 C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 22-Jun-1999
 C;Accession: S55205; S57031; S43744
 R;de Haan, M.; Smits, P.H.M.; Grivell, L.A.
 submitted to the EMBL Data Library, May 1995
 A;Reference number: S55183
 A;Accession: S55205
 A;Molecule type: DNA
 A;Residues: 1-585 <DEH>
 A;Cross-references: EMBL:X87611; NID:9854567; PIDN:CAA60939.1; PID:9854590
 R;de Haan, M.; Grivell, L.A.; Smits, P.H.M.
 submitted to the Protein Sequence Database, September 1995
 A;Reference number: S56771
 A;Accession: S57031
 A;Molecule type: DNA
 A;Residues: 1-585 <ZAG>
 A;Cross-references: EMBL:X49516; NID:91015649; PIDN:CAA89540.1; PID:91015650; MIPS:YJR01
 R;Velasco, J.A.; Canedo, J.; Pena, M.C.; Kawakami, T.; Laborda, J.; Notario, V.
 submitted to the EMBL Data Library, January 1994
 A;Description: Cloning of the dihydroxyacid dehydratase gene (ILV-3) from the yeast Sacc
 A;Reference number: S43744
 A;Accession: S43744
 A;Molecule type: DNA
 A;Residues: 'MF', 84-85, 'TEKR', 90-237, 'S', 239-241, 'S', 243-320, 'P', 322-452 <VEL>
 A;Cross-references: EMBL:L13975; NID:9439111; PIDN:AAA34568.1; PID:9553125
 A;Genetics:
 A;Gene: SGD:ILV3
 A;Cross-references: SGD:S0003777; MIPS:YJR016c
 A;Map position: 10R
 C;Superfamily: dihydroxy-acid dehydratase
 C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 1.1%; Score 7; DB 2; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 ASNGCVL 241
 Db 577 ASNGCVL 583

RESULT 24

T17333
 hypothetical protein DKFP434E15.1 - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T17333
 R;Dierkerhoef, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, September 1999
 A;Reference number: Z18727
 A;Accession: T17333
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-596 <DUE>
 A;Cross-references: EMBL:AL117632
 A;Experimental source: adult testis; clone DKFP434E15

C:Geneid: 8
A:Note: DKFp434E165.1

Query Match
Best Local Similarity 1.1%; Score 7; DB 2; Length 596;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 549 STPLSTS 555
|||||
DB 388 STPLSTS 394

RESULT 25
T35054
probable transport system permease protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35054
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z21566
A:Accession: T35054
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-601 <SEB>
A:Cross-references: EMBL:AL031371; PIDN:CAA20556.1; GSPDB:GN00070; SCOEDB:SC4G2.19
C:Geneid: 8
A:Gene: SCOEDB:SC4G2.19

Query Match
Best Local Similarity 1.1%; Score 7; DB 2; Length 601;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 AACQRPV 355
|||||
DB 582 AACQRPV 588

RESULT 26
G90575
transketolase (tk) [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: G90575
R:Shamdan, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A93512; MUID:21267165; PMID:11353084
A:Accession: G90575
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-615 <KUR>
A:Cross-references: GB:AL445566; PID:G1408925; PIDN:CAC13684.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Geneid: 8
A:Gene: MYPV 5110
A:Geneid: 8
A:Superfamily: transketolase; thiamin pyrophosphate-binding domain homology

Query Match
Best Local Similarity 1.1%; Score 7; DB 2; Length 615;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 IEKAKS 236
|||||
DB 222 IEKAKS 228

RESULT 27
B86431
hypothetical protein T518.4 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: B86431
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, W.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroe, J.S.; Malt, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86431
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-622 <STO>
A:Cross-references: GB:AE005172; NID:G4587515; PIDN:AAD25746.1; GSPDB:GN00141
C:Geneid: 8
A:Map position: 1

Query Match
Best Local Similarity 1.1%; Score 7; DB 2; Length 622;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 ISPNFNF 287
|||||
DB 537 ISPNFNF 543

RESULT 28
A39262
transcription factor HNF-1A - mouse
N:Alternate names: hepatic nuclear factor 1-alpha; transcription factor APF; transcript:
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A39262; S70436
R:Kuo, C.T.; Conley, P.B.; Hsieh, C.L.; Franke, U.; Crabtree, G.R.
Proc. Natl. Acad. Sci. U.S.A. 87, 9838-9842, 1990
A:Title: Molecular cloning, functional expression, and chromosomal localization of mouse
A:Reference number: A39262; MUID:91088607; PMID:2263635
A:Accession: A39262
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-628 <KUC>
A:Cross-references: GB:M57966; NID:G193885; PIDN:AAA37821.1; PID:G193886
R:Bach, I.; Pontoglio, M.; Yaniv, M.
Nucleic Acids Res. 20, 4199-4204, 1992
A:Title: Structure of the gene encoding hepatocyte nuclear factor 1 (HNF1).
A:Reference number: S70435; MUID:92375726; PMID:1354855
A:Accession: S70436
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 199-279 <BAC>
A:Note: only a part of the nucleic acid sequence is shown
C:Geneid: 8
A:Gene: Hnf-1
A:Introns: 238/2
A:Note: the list of introns is incomplete
C:Complex: homodimer; can also form heterodimers with, for example, HNF-1B
C:Function:
A:Description: transcription activator required for the expression of a number of liver.
A:Note: also expressed in other tissues (kidney, stomach, intestine), where it may play
C:Superfamily: transcription factor HNF-1; homeobox homology
C:Keywords: DNA binding; heterodimer; homeobox; homodimer; liver; nucleus; transcription
F:1-33/Region: dimerization
F:222-277/Domain: homeobox homology <HGX>

Query Match
Best Local Similarity 1.1%; Score 7; DB 1; Length 628;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 344 SATSENA 350

Db 325 SATSEAA 331

RESULT 29

transcription factor HNF-1A - rat
 N:Alternate names: hepatic nuclear factor 1-alpha; transcription factor APF; transcript
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A33333; A34590; S11568; S70435
 R:Frain, M.; Swart, G.; Monaci, P.; Nicotia, A.; Staemfli, S.; Frank, R.; Cortese, R.
 Cell 59, 145-157, 1989
 A>Title: The liver-specific transcription factor LF-B1 contains a highly diverged homeo
 A:Reference number: A33333; MUID:90003224; PMID:2571419
 A:Accession: A33333
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-628 <PRA>
 A:Cross-references: GB:J03170; NID:9205164; PIDN:AAA1524.1; PID:9205165
 R:Baumhueter, S.; Mendel, D.B.; Conley, P.B.; Kuo, C.J.; Turk, C.; Graves, M.K.; Edwards
 Genes Dev. 4, 372-379, 1990
 A>Title: HNF-1 shares three sequence motifs with the POU domain proteins and is identica
 A:Reference number: A34590; MUID:90249741; PMID:1970973
 A:Accession: A34590
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 'R', 166-628 <BAU>
 A:Cross-references: GB:X53297; NID:957867; PIDN:CAA37387.1; PID:957868
 A:Note: the authors translated the codon GAG for residue 616 as Asp
 R:Chouard, T.; Blumenfeld, M.; Bach, I.; Vandekerckhove, J.; Cerejint, S.; Yaniv, M.
 Nucleic Acids Res. 18, 5853-5863, 1990
 A>Title: A distal dimerization domain is essential for DNA-binding by the atypical HNF1
 A:Reference number: S11568; MUID:91016926; PMID:2216777
 A:Accession: S11568
 A:Molecule type: mRNA
 A:Residues: 1-628 <CHO>
 A:Cross-references: EMBL:X54423; NID:956367; PIDN:CAA38295.1; PID:956368
 R:Bach, I.; Fontoglio, M.; Yaniv, M.
 Nucleic Acids Res. 20, 4199-4204, 1992
 A>Title: Structure of the gene encoding hepatocyte nuclear factor 1 (HNF1).
 A:Reference number: S70435; MUID:92375726; PMID:1354855
 A:Accession: S70435
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 199-279 <BAC>
 A:Note: only a part of the nucleic acid sequence is shown
 C:Genetics:
 A:Gene: Hnf-1
 A:Introns: 238/2
 A:Note: the list of introns is incomplete
 C:Complex: homodimer; can also form heterodimers with, for example, HNF-1B
 C:Function:
 A:Description: transcription activator required for the expression of a number of liver-
 A:Note: also expressed in other tissues (kidney, stomach, intestine), where it may play
 C:Superfamily: transcription factor HNF-1; homeobox homology
 C:Keywords: DNA binding; heterodimer; homeobox; homodimer; liver; nucleus; transcription
 F:1-33/Region: dimerization
 F:222-277/Domain: homeobox homology <HOK>

Query Match 1.1%; Score 7; DB 1; Length 628;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 344 SATSEAA 350
 Db 325 SATSEAA 331

RESULT 30
 G90779
 hypothetical protein Eca1207 [imported] - Escherichia coli (strain O157:H7, substrain RI
 C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: G90779
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
 A:Reference number: A9629; MUID:21156231; PMID:11258796
 A:Accession: G90779
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-634 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA834630.1; PID:913360667; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: Eca1207

Query Match 1.1%; Score 7; DB 2; Length 634;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 417 LHGFSS 423
 Db 617 LHGFSS 623

RESULT 31

hypothetical protein - human herpesvirus 4
 C:Species: human herpesvirus 4, Epstein-Barr virus
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Jan-1995
 C:Accession: S32991
 R:Farrell, P.J.
 submitted to the EMBL Data Library, March 1988
 A:Reference number: S32973
 A:Accession: S32991
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-657 <FAR>
 A:Cross-references: EMBL:V01555

Query Match 1.1%; Score 7; DB 2; Length 657;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 324 PAVSEGG 330
 Db 379 PAVSEGG 385

RESULT 32

F81896
 probable phage transposase NMA1284 [imported] - Neisseria meningitidis (strain Z2491 ser
 C:Species: Neisseria meningitidis
 C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C:Accession: F81896
 R:Farhail, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A:Reference number: A81775; MUID:20222556; PMID:10761919
 A:Accession: F81896
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-657 <PAR>
 A:Cross-references: GB:AL162755; GB:AL157959; NID:97379742; PIDN:CAB84536.1; PID:9737996
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: NMA1284

Query Match 1.1%; Score 7; DB 2; Length 657;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 TERLVAL 18
|||||
Db 178 TERLVAL 184

RESULT 33
A81125
bacteriophage transposase NMB1081 [imported] - Neisseria meningitidis (strain MCS8 serog
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C/Accession: A81125
R/RetelIn, H.; Sanders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A/Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MCS8.
A/Reference number: A81000; MUID:20175755; PMID:10710307
A/Accession: A81125
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-661 <VER>
A/Cross-references: GB:AE002458; GB:AE002098; NID:g7226311; PIDN:AAF41473.1; PID:g722631
A/Experimental source: serogroup B, strain MCS8
C/Genetics:
A/Gene: NMB1081

Query Match 1.1%; Score 7; DB 2; Length 661;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 TERLVAL 18
|||||
Db 182 TERLVAL 188

RESULT 34
S45131
probable membrane protein YNR008w - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein N2042
C/Species: Saccharomyces cerevisiae
C/Date: 03-May-1994 #sequence_revision 02-Aug-1994 #text_change 19-Apr-2002
R/Verhasselt, P.; Aert, R.; Voet, M.; Volckaert, G.
A/Accession: S45131; S48345; S63334
Submitted to the EMBL Data Library, January 1994
A/Description: Twelve open reading frames revealed on the 23.6 kbp segment flanking the
A/Reference number: S45119
A/Accession: S45131
A/Molecule type: DNA
A/Residues: 1-661 <VER>
A/Cross-references: EMBL:X77395; NID:g496717; PID:g496725
R/Verhasselt, P.; Aert, R.; Voet, M.; Volckaert, G.
Yeast 10, 1355-1361, 1994
A/Title: Twelve open reading frames revealed in the 23.6 kb segment flanking the centrom
A/Reference number: S48338; MUID:95208356; PMID:7900425
A/Accession: S49345
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-661 <VE2>
A/Cross-references: EMBL:X77395; NID:g496717; PIDN:CAA54576.1; PID:g496725
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1994
R/Aert, R.; Verhasselt, P.; Voet, M.; Volckaert, G.
Submitted to the Protein Sequence Database, April 1996
A/Reference number: S62910
A/Accession: S63334
A/Molecule type: DNA
A/Residues: 1-661 <AER>
A/Cross-references: EMBL:Z71623; NID:g1302481; PID:e239782; PID:g1302482; MIPS:YNR008w
A/Experimental source: strain S288C
C/Genetics:
A/Gene: SGD:LR01
A/Cross-references: SGD:S0005291
A/Map position: 14R

C/Keywords: transmembrane protein
F:80-96/Domain: transmembrane #status predicted <TM>

Query Match 1.1%; Score 7; DB 2; Length 661;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 SSSEPAL 427
|||||
Db 431 SSSEPAL 437

RESULT 35
C81505
hypothetical protein CP1075 [imported] - Chlamydomonas reinhardtii (strain AR39)
C/Species: Chlamydomonas reinhardtii, Chlamydia pneumoniae
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C/Accession: C81505
R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salberg
Nucleic Acids Res. 28, 1397-1406, 2000
A/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A/Reference number: A81500; MUID:20150255; PMID:10684935
A/Accession: C81505
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-674 <REA>
A/Cross-references: GB:AE002264; GB:AE002161; NID:g7189984; PIDN:AAF38847.1; PID:g718998
A/Experimental source: strain AR39, HL cells
C/Genetics:
A/Gene: CP1075

Query Match 1.1%; Score 7; DB 2; Length 674;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 LIQHSAX 74
|||||
Db 414 LIQHSAX 420

RESULT 36
E72033
hypothetical protein - Chlamydomonas reinhardtii (strain CWL029)
C/Species: Chlamydomonas reinhardtii, Chlamydia pneumoniae
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C/Accession: E72033
R/Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A/Reference number: A72000; MUID:9206606; PMID:10192388
A/Accession: E72033
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-680 <ARN>
A/Cross-references: GB:AE001661; GB:AE001363; NID:g4377104; PIDN:AAD18934.1; PID:g437710
A/Experimental source: strain CWL029
C/Genetics:
A/Gene: CPN0796

Query Match 1.1%; Score 7; DB 2; Length 680;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 LIQHSAX 74
|||||
Db 420 LIQHSAX 426

RESULT 37
B86590
hypothetical protein CPJ0796 [imported] - Chlamydomonas reinhardtii (strain J138)
C/Species: Chlamydomonas reinhardtii, Chlamydia pneumoniae

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C/Accession: B86590
 R/Shiira, M.; Hirakawa, H.; Kinoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
 Nucleic Acids Res. 28, 2311-2314, 2000
 A>Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A/Reference number: A86491; MUID:20330349; PMID:10871362
 A/Accession: B86590
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-680 <STO>
 A/Cross-references: GB:BA000008; NID:g8979169; PIDN:BAA99004.1; GSPDB:GN00142
 A/Experimental source: strain J138
 C/Genetics:
 A/Gene: Cpj0796

Query Match 1.1%; Score 7; DB 2; Length 680;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 LIQHSK 74
 |||||
 Db 420 LIQHSK 426

RESULT 38
 G82618
 plus biogenesis protein Xp1953 [imported] - Xylella fastidiosa (strain 9asc)
 C/Species: Xylella fastidiosa
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C/Accession: G82618
 R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A/Reference number: A82515; MUID:20355717; PMID:10910347
 A/Note: for a complete list of authors see reference number A59328 below
 A/Accession: G82618
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-693 <SIM>

A/Cross-references: GB:AE004014; GB:AE003849; NID:g9107044; PIDN:AAF84755.1; GSPDB:GN001
 A/Experimental source: strain 9asc
 R/Simpson, A.J.G.; Reinach, F.C.; Arrudá, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Britone, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carreir, H
 A/Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000
 A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigz
 Chado, M.A.; Madalita, A.M.B.N.; Madalita, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B
 A/uthors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sasaki
 A/uthors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tuhaco, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A/Reference number: A59328
 A/Contents: annotation
 C/Genetics:
 A/Gene: XF1953

Query Match 1.1%; Score 7; DB 2; Length 693;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 390 LEDSNKL 396
 |||||
 Db 295 LEDSNKL 301

RESULT 39
 S01991
 transforming protein B-myb - human
 C/Species: Homo sapiens (man)
 C/Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 18-Jun-1999
 C/Accession: S01991; S48661

R/Nomura, N.; Takahashi, M.; Matsui, M.; Ishii, S.; Date, T.; Saeamoto, S.; Ishizaki, R
 Nucleic Acids Res. 16, 11075-11089, 1988
 A>Title: Isolation of human cDNA clones of myb-related genes, A-myb and B-myb.
 A/Reference number: S01991; MUID:89083548; PMID:3060855
 A/Accession: S01991
 A/Molecule type: mRNA
 A/Residues: 1-700 <NM>
 A/Cross-references: EMBL:X13293; NID:g29471; PIDN:CAA31655.1; PID:g29472
 R/Takemoto, Y.; Tashiro, S.; Handa, H.; Ishii, S.
 FEBS Lett. 350, 55-60, 1994
 A>Title: Multiple nuclear localization signals of the B-myb gene product.
 A/Reference number: S48661; MUID:94341374; PMID:8062924
 A/Contents: annotation; identification of nuclear location signal regions by mutational
 C/Genetics:
 A/Gene: GDB:MYB12; BMTB
 A/Cross-references: GDB:128109; OMIM:310305; OMIM:601415
 A/Map position: Xq13-Xq13
 C/Function:
 A/Description: transcription regulation; widespread activator of cell cycle genes; repre
 C/Superfamily: myb transforming protein; myb DNA-binding repeat homology
 C/Keywords: DNA binding; duplication; nucleus; transcription regulation
 F/26-77/Domain: myb DNA-binding repeat homology <MYB1>
 F/78-129/Domain: myb DNA-binding repeat homology <MYB2>
 F/130-180/Domain: myb DNA-binding repeat homology <MYB3>
 F/411-417/Region: nuclear location signal
 F/569-584/Region: nuclear location signal

Query Match 1.1%; Score 7; DB 1; Length 700;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 CSQKVV 87
 |||||
 Db 479 CSQKVV 485

RESULT 40
 S33704
 transforming protein B-myb - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
 C/Accession: S33704; 148269
 R/Lam, E.W.F.; Robinson, C.; Watson, R.J.
 Oncogene 7, 1885-1890, 1992
 A>Title: Characterization and cell cycle-regulated expression of mouse B-myb.
 A/Reference number: S33704; MUID:92366176; PMID:1501895
 A/Accession: S33704
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-704 <LAM>
 A/Cross-references: EMBL:X70472; NID:g312825; PIDN:CAA49898.1; PID:g312826
 R/Lam, E.W.; Watson, R.J.
 EMBO J. 12, 2705-2713, 1993
 A>Title: An E2F-binding site mediates cell-cycle regulated repression of mouse B-myb tra
 A/Reference number: 148269; MUID:93327760; PMID:8334969
 A/Accession: 148269
 A/Status: preliminary; translated from GB/EMBL/DDB3
 A/Molecule type: DNA
 A/Residues: 1-6 <RES>
 A/Cross-references: EMBL:X73028; NID:g297153; PIDN:CAA51511.1; PID:g581921
 C/Genetics:
 A/Gene: B-myb
 C/Function:
 A/Description: transcription regulation; widespread activator of cell cycle genes; repre
 C/Superfamily: myb transforming protein; myb DNA-binding repeat homology
 C/Keywords: DNA binding; duplication; nucleus; transcription regulation
 F/26-77/Domain: myb DNA-binding repeat homology <MYB1>
 F/78-129/Domain: myb DNA-binding repeat homology <MYB2>
 F/130-180/Domain: myb DNA-binding repeat homology <MYB3>
 F/414-420/Region: nuclear location signal
 F/573-588/Region: nuclear location signal

Query Match 1.1%; Score 7; DB 1; Length 704;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 CSQKVV 87
DB 482 CSQKVV 488

RESULT 41

T38117
probable protein kinase - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C/Accession: T38117

R/Conor, R.; Church, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1996

A/Reference number: Z21771

A/Accession: T38117

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1704 <CON>

A/Cross-references: EMBL:Z98600; PIDN:CA81253.1; GSPDB:GN00066; SPDB:SPAC20G4.03C

A/Experimental source: strain 972h-; cosmid c20G4

C/Genetics:

A/Gene: SPDB:SPAC20G4.03C

A/Map position: 1

Query Match 1.1%; Score 7; DB 2; Length 704;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LVALLS 21
DB 50 LVALLS 56

RESULT 42

T17002
probable beta-galactosidase (EC 3.2.1.23) precursor - apple tree

C/Species: Malus domestica (apple tree)

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000

C/Accession: T17002

R/Rosa, G.S.; Megryzn, T.; Macrae, E.A.; Redgwell, R.J.

Plant Physiol. 106, 521-528, 1994

A/Title: Apple beta-galactosidase. Activity against cell wall polysaccharides and charac

A/Reference number: Z18645; MUID:95083752; PMID:7991682

A/Accession: T17002

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: mRNA

A/Residues: 1731 <ROS>

A/Cross-references: EMBL:L29451; NID:g507277; PIDN:AAA62324.1; PID:g507278

A/Experimental source: strain Granny Smith; cortical

C/Superfamily: beta-galactosidase bga

C/Keywords: glycosidase; hydrolase

F/1-23/Domain: signal sequence #status predicted <SIG>

F/24-731/Product: probable beta-galactosidase #status predicted <MAT>

Query Match 1.1%; Score 7; DB 2; Length 731;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 648 KVGSSS 654
DB 414 KVGSSS 420

RESULT 43

C99995
hypothetical protein Ec82931 [imported] - Escherichia coli (strain O157:H7, substrain RI

C/Species: Escherichia coli

C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002

C/Accession: C99995

R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001

A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A/Reference number: A99629; MUID:21156231; PMID:11238796

A/Accession: C99995

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1745 <HAY>

A/Cross-references: GB:BA000007; PIDN:BA836354.1; PID:g13362400; GSPDB:GN00154

A/Experimental source: strain O157:H7, substrain RIMD 0509952

C/Genetics:

A/Gene: Ec82931

C/Superfamily: Escherichia coli hypothetical 68.5K protein (moIR-DglX region)

Query Match 1.1%; Score 7; DB 2; Length 745;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 417 LHGFSS 423
DB 728 LHGFSS 734

RESULT 44

H85840
hypothetical protein yehQ [imported] - Escherichia coli (strain O157:H7, substrain EDL93

C/Species: Escherichia coli

C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 03-Jun-2002

C/Accession: H85840

R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grobeck, E.J.; Davis, N.W.; Llm, A.; Dimalanta, E.; Peticola, K.; Apodaca,

Nature 409, 529-533, 2001

A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A/Reference number: A85480; MUID:21074935; PMID:11206551

A/Accession: H85840

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1745 <STO>

A/Cross-references: GB:AE005174; NID:g12516346; PIDN:AA657188.1; GSPDB:GN00145; UNGP:23

A/Experimental source: strain O157:H7, substrain EDL933

C/Genetics:

A/Gene: yehQ

C/Superfamily: Escherichia coli hypothetical 68.5K protein (moIR-DglX region)

Query Match 1.1%; Score 7; DB 2; Length 745;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 417 LHGFSS 423
DB 728 LHGFSS 734

RESULT 45

E82013
probable transcription accessory protein MMA0194 [imported] - Neisseria meningitidis (str

C/Species: Neisseria meningitidis

C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C/Accession: E82013

R/Fairchild, J.; Achman, M.; James, K.D.; Bentley, S.D.; Church, C.; Klee, S.R.; Morel

i; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A/Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A/Reference number: AB1775; MUID:20222556; PMID:10761919

A/Accession: E82013

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1757 <PAR>

A/Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CA83508.1; PID:g737896

A/Experimental source: serogroup A, strain Z2491

C/Genetics:

A/Gene: MMA0191; MMA0194

C/Superfamily: hypothetical protein ydcI

C:Keywords: methyltransferase; S-adenosylmethionine

Query Match 1.1%; Score 7; DB 2; Length 757;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 638 SENSRSRE 644
|||||
DB 723 SENSRSRE 729

RESULT 46

A81242

transcription accessory protein Tex, probable NMB0075 [imported] - Neisseria meningitidis

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: A81242

R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

Science 287, 1809-1815, 2000

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MCS8.

A:Reference number: A81000; MUID:2015755; PMID:10710307

A:Accession: A81242

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-757 <TET>

A:Cross-references: GB:AE002366; GB:AE002098; NID:g7225284; PIDN:AAF40542.1; PID:g722529

A:Experimental source: serogroup B, strain MCS8

C:Genetics:

A:Gene: NMB0075

C:Superfamily: hypothetical protein ydc1

Query Match 1.1%; Score 7; DB 2; Length 757;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 638 SENSRSRE 644
|||||
DB 723 SENSRSRE 729

RESULT 47

E97224

ATP-dependent Lon protease [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 02-Jun-2003

C:Accession: E97224

R:Nolling, U.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: E97224

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-778 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK80584.1; PID:g15025664; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC2637

C:Superfamily: ATP-dependent Lon protease

Query Match 1.1%; Score 7; DB 2; Length 778;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 659 MEIIEVS 665
|||||
DB 484 MEIIEVS 490

RESULT 48

T51759

[glutamate-ammmonia-lyase] adenylyltransferase (EC 2.7.7.42) [imported] - Streptomyces c

C:Species: Streptomyces coelicolor

C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 03-Jun-2002

C:Accession: T51759

R:Pink, D.; Fialke, D.; Wohlleben, W.; Engels, A.

Microbiology 145, 2313-2322, 1999

A:Title: Nitrogen metabolism in Streptomyces coelicolor A3(2): modification of glutamine

A:Reference number: Z25448

A:Accession: T51759

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-784 <FIN>

A:Cross-references: EMBL:Y17736; PIDN:CAA76840.2

C:Genetics:

A:Gene: glnB

C:Keywords: nucleotidyltransferase

Query Match 1.1%; Score 7; DB 2; Length 784;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 IDEAYRF 273
|||||
DB 410 IDEAYRF 416

RESULT 49

T34468

hypothetical protein ZK770.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34468

R:Maggi, L.; Gattung, S.; Bartko, L.

submitted to the EMBL Data Library, April 1997

A:Description: The sequence of C. elegans cosmid ZK770.

A:Reference number: Z21530

A:Accession: T34468

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-795 <MAG>

A:Cross-references: EMBL:U97404; PIDN:AA893309.1; GSPDB:GN00019; CESP:ZK770.1

A:Experimental source: strain Bristol N2; clone ZK770

C:Genetics:

A:Gene: CESP:ZK770.1

A:Map position: 1

A:introns: 18/1; 84/2; 125/3; 198/3; 257/3; 337/3; 378/3; 424/3; 477/2; 510/1; 563/3; 63

Query Match 1.1%; Score 7; DB 2; Length 795;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 406 SVSYSAS 412
|||||
DB 665 SVSYSAS 671

RESULT 50

T03544

hypothetical protein - Rhodobacter capsulatus

C:Species: Rhodobacter capsulatus

C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999

C:Accession: T03544

R:Vlcek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.

Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997

A:Title: Sequence of a 189-bp segment of the chromosome of Rhodobacter capsulatus SB1003

A:Reference number: Z14955; MUID:97404404; PMID:9256491

A:Accession: T03544

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-828 <VLC>

A:Cross-references: EMBL:AF010496; NID:g3128256; PIDN:AAI6197.1; PID:g3128345

C:Genetics:
A:Map position: 1

Query Match 1.1%; Score 7; DB 2; Length 828;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 119 HLAGGF 125
|||
Db 666 HLAGGF 672

Search completed: June 21, 2004, 13:28:58
Job time : 25 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:22:43 ; Search time 17 Seconds

(without alignments)
2036.862 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665
Sequence: 1 MAHMGICIVTERLVALL.....LKVSGSSFSGSMETIEVS 665Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 33406

Minimum DB seq length: 473

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	71.0	665	1 DUSG_HUMAN	Q9bY84 homo sapien
2	20	3.0	625	1 DUS8_HUMAN	Q13202 homo sapien
3	20	3.0	663	1 DUS8_MOUSE	009112 mus musculu
4	9	1.4	619	1 VHP1_CAEEL	010038 caenorhabdi
5	7	1.1	473	1 ATPB_PTRES	003080 pteridium e
6	7	1.1	478	1 LXC2_PROLE	P29236 photobacter
7	7	1.1	507	1 MS14_ARATH	O22607 arabidopsi
8	7	1.1	510	1 PUR1_CHICK	P28173 gallus gall
9	7	1.1	524	1 HEMK_RICCN	Q92913 rickettsia
10	7	1.1	535	1 TCER1_AYESA	P40412 avena sativ
11	7	1.1	535	1 TCER2_AYESA	P54411 avena sativ
12	7	1.1	535	1 TCPE_ARATH	Q04450 arabidopsi
13	7	1.1	564	1 NOX1_HUMAN	Q9y588 homo sapien
14	7	1.1	566	1 SCC1_YEAST	Q12158 saccharomyc
15	7	1.1	579	1 MALT1_ARDAE	P13080 aedes aegypt
16	7	1.1	583	1 SCF3_SCHPO	P41000 schizosacch
17	7	1.1	585	1 ILV3_YEAST	P39522 saccharomyc
18	7	1.1	591	1 UL49_EBV	P14347 Epstein-Barr
19	7	1.1	597	1 Y745_ARATH	P59278 arabidopsi
20	7	1.1	628	1 HNFA_MOUSE	P22861 mus musculu
21	7	1.1	628	1 HNFA_MOUSE	P15557 ratu
22	7	1.1	661	1 PDAT_YEAST	P40345 saccharomyc
23	7	1.1	668	1 V244_POMPV	Q91426 fowlipox vir
24	7	1.1	700	1 MYBB_HUMAN	P10244 homo sapien
25	7	1.1	704	1 MYBB_HUMAN	P48912 mus musculu
26	7	1.1	731	1 BGAL_MALDO	P48911 malus domes
27	7	1.1	757	1 YHGF_NEIMA	P57072 neisseria m
28	7	1.1	757	1 YHGF_NEIMA	Q51152 neisseria m
29	7	1.1	795	1 DEG3_CAEEL	001635 caenorhabdi
30	7	1.1	843	1 NAL6_MOUSE	Q91982 mus musculu
31	7	1.1	854	1 NAL6_MOUSE	Q63035 ratu
32	7	1.1	892	1 NAL6_MOUSE	P59044 homo sapien
33	7	1.1	938	1 TRG1_ECOLI	P33790 escherichia

34	7	1.1	945	1 AMPE_RAT	P50123 ratu
35	7	1.1	966	1 RHG4_HUMAN	P98171 homo sapien
36	7	1.1	966	1 MMBB_MYCTU	O53653 mycobacteri
37	7	1.1	1038	1 KF17_MOUSE	Q99PW8 mus musculu
38	7	1.1	1071	1 UBPT_YEAST	P40453 saccharomyc
39	7	1.1	1099	1 CVAT_MOUSE	P51829 mus musculu
40	7	1.1	1163	1 LEPR_MACMU	Q9mY10 macaca mula
41	7	1.1	1165	1 LEPR_HUMAN	P48357 homo sapien
42	7	1.1	1168	1 DHX8_ARATH	Q38953 arabidopsi
43	7	1.1	1275	1 AIP1_MOUSE	Q9wvq1 mus musculu
44	7	1.1	1277	1 AIP1_RAT	Q9wvq1 mus musculu
45	7	1.1	1318	1 YAB2_SCHPO	Q09804 schizosacch
46	7	1.1	1332	1 YS1_CAEEL	Q20255 caenorhabdi
47	7	1.1	1411	1 TCOF_HUMAN	Q13428 homo sapien
48	7	1.1	1455	1 AIP1_HUMAN	Q86118 homo sapien
49	7	1.1	1472	1 MYHB_HUMAN	P35749 homo sapien
50	7	1.1	2077	1 TEGU_HSVG	P30002 human herpe
51	7	1.1	2077	1 TEGU_HSVG	P52340 human herpe
52	7	1.1	2210	1 RRPO_LYCV	P14240 lymphocytic
53	7	1.1	2611	1 BPIE_MOUSE	Q91208 mus musculu
54	7	1.1	2768	1 THYG_HUMAN	P01266 homo sapien
55	7	1.1	3060	1 BPEB_HUMAN	Q8wXK8 homo sapien
56	7	1.1	3214	1 BPA1_HUMAN	Q03001 homo sapien
57	7	1.1	5038	1 PCLO_MOUSE	Q9gyx7 mus musculu
58	7	1.1	5085	1 PCLO_RAT	Q9jK66 ratu
59	7	1.1	5147	1 PCLO_HUMAN	Q9yCv0 homo sapien
60	7	1.1	5171	1 BPEA_HUMAN	Q94833 homo sapien
61	7	1.1	5262	1 MLR2_HUMAN	Q14686 homo sapien
62	7	1.1	7389	1 BPA1_MOUSE	Q91206 mus musculu
63	6	0.9	473	1 ATPB_BACSU	P37809 bacillus su
64	6	0.9	473	1 C1SY_SCHPO	Q10306 schizosacch
65	6	0.9	473	1 IDH_ANASP	P50214 anabaena sp
66	6	0.9	473	1 V12_HPV16	P03107 human papil
67	6	0.9	473	1 ZF38_HUMAN	Q9y5a6 homo sapien
68	6	0.9	475	1 DLDH_BUCBP	Q896q8 buchnera ap
69	6	0.9	475	1 PAAB_ECOLI	P76083 escherichia
70	6	0.9	475	1 RMIC_ECOLI	P27850 escherichia
71	6	0.9	478	1 SEB4_HUMAN	O43236 homo sapien
72	6	0.9	478	1 SEB4_MOUSE	P28641 mus musculu
73	6	0.9	479	1 YP66_YEAST	Q12194 saccharomyc
74	6	0.9	480	1 UCRI_BOVIN	P31800 bos taurus
75	6	0.9	481	1 MENE_BACAA	Q81x97 bacillus an

ALIGNMENTS

RESULT 1

ID DUSG_HUMAN STANDARD; PRT; 665 AA.

AC Q9BY84; Q9C0G3; 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Dual specificity protein phosphatase 16 (EC 3.1.3.48) (EC 3.1.3.16)

DE (Mitogen-activated protein kinase phosphatase 7) (MAP kinase

DE phosphatase 7) (MKP-7).

DE DUSP16 OR MKP7 OR KIAA1700.

GN Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_Taxid=9606;

RN [1] SEQUENCE FROM N.A.

RX MEDLINE=21486429; PubMed=11489891;

RA Masuda K., Shima H., Watanabe M., Kikuchi K.;

RT "MKP-7, a novel mitogen-activated protein kinase phosphatase,

RT functions as a shuttle protein.";

RL J. Biol. Chem. 276:39002-39011(2001).

RN [2] SEQUENCE FROM N.A.

RP TISSUE=Brain;

RX MEDLINE=21082932; PubMed=11214970;

RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIX.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:347-355(2000).
 CC -1- FUNCTION: Involved in the inactivation of MAP kinases.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; AB052156; BAB40814.1; -;
 DR EMBL; AB051487; BAB2191.1; ALT_INTT.
 DR HSSP; O16828; IMKP.
 DR Genew; HGNC:17909; DUSP16.
 DR MIM; 607175; -;
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0004721; F:protein phosphatase activity; TAS.
 DR GO; GO:0016311; F:dephosphorylation; TAS.
 DR GO; GO:0000188; P:inactivation of MAPK; TAS.
 DR GO; GO:0045209; P:leptomycin B-sensitive MAPK phosphatase nuc. .; TAS.
 DR GO; GO:0045204; P:MAPK nucleus export; TAS.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF00581; Rhodanese; 1.
 DR PRINTS; PR01764; MAPKPHPTASB.
 DR SMART; SM00195; DSPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS50206; RHODANES_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KM Hydroxylase; Nuclear protein.
 FT DOMAIN 22 137 RHODANES.
 FT DOMAIN 227 289 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 244 244 PHOSPHOCYSTEINE INTERMEDIATE (BY
 FT ACT_SITE 244 244 SIMILARITY)
 SQ SEQUENCE 665 AA; 73101 MW; 1BD853FF0846DDFF CRC64;
 Query Match 71.0%; Score 472; DB 1; Length 665;
 Best Local Similarity 100.0%; Pred. 0;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 374 SPLVALSGHLMSADRLSDSNLKRFSFLDIKSYSSMASLHGSSSEDALFYKPS 433
 QY 434 TLLDGTNNLCQPSVQVSEQTPETSPPKREASIPKLIQTARPSDSQSKRLHVSSTSSG 493
 Db 434 TLLDGTNNLCQPSVQVSEQTPETSPPKREASIPKLIQTARPSDSQSKRLHVSSTSSG 493
 QY 494 TQKRLSLPLHRSSGVEDNYHTSPFGISTGQOHLTKSAGLGKWMHSDIILAPQTSPL 553
 Db 494 TQKRLSLPLHRSSGVEDNYHTSPFGISTGQOHLTKSAGLGKWMHSDIILAPQTSPL 553
 QY 554 TSSWYFATESSHFYASAIYGGASAYSCSQPTCGDQYVYRRQKPSDRADSRSSW 613
 Db 554 TSSWYFATESSHFYASAIYGGASAYSCSQPTCGDQYVYRRQKPSDRADSRSSW 613
 QY 614 HEESPFERQFKRRSCQMEFGESIMSENRSRELKVGSSQSSFSQSMELIEVS 665
 Db 614 HEESPFERQFKRRSCQMEFGESIMSENRSRELKVGSSQSSFSQSMELIEVS 665
 RESULT 2
 ID DUS8_HUMAN STANDARD; PRT; 625 AA.
 AC Q13202;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Dual specificity protein phosphatase hvh-5).
 GN DUSP8 OR VH5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=96009533; PubMed=7561881;
 RA Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.;
 RT "hvh-5, a protein tyrosine phosphatase abundant in brain that
 RT inactivates mitogen-activated protein kinase.";
 RL J. Neurochem. 65:1823-1833(1995).
 CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
 CC phosphate as well as with serine/threonine-protein phosphate (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- TISSUE SPECIFICITY: Abundant in brain, heart and skeletal muscle.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Non-receptor class dual specificity subfamily.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; U27193; AAA83151.1; -;
 DR HSSP; O16828; IMKP.
 DR Genew; HGNC:3074; DUSP8.
 DR MIM; 602038; -;
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0004725; F:protein tyrosine phosphatase activity; TAS.
 DR GO; GO:0000188; P:inactivation of MAPK; TAS.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR008343; MAPK_phosph.
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR_phosphatase.

```

DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolyase; Nuclear protein.
KM DOMAIN 23 138 RHODANSE.
FT DOMAIN 162 430 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 310 550 PRO-RICH.
FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
SQ SEQUENCE 625 AA; 65840 MW; DCBEA14487219666 CRC64;

Query Match 3.0%; Score 20; DB 1; Length 625;
Best Local Similarity 100.0%; Pred. No. 6.4e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 VHCLAGISRSATIAIAYIMK 261
Db 244 VHCLAGISRSATIAIAYIMK 263

RESULT 3
DUS8_MOUSE STANDARD; PRT; 663 AA.
ID DUS8_MOUSE 009112;
AC 009112;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Neuronal tyrosine threonine phosphatase 1).
DE DUSP8 OR NTP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
ON NCBI_TaxID=10090;
RX MEDLINE=69311565; PubMed=873137;
RA Theodosiou A.M., Rodrigues N.R., Nesbit M.A., Ambrose H.J.,
RA Paterson H., McEllan-Arnold E., Boyd Y., Levertha M.A., Owen N.,
RA Blake D.J., Ashworth A., Davies K.E.;
RT "A member of the MAP kinase phosphatase gene family in mouse
RT containing a complex trinucleotide repeat in the coding region.";
RL Hum. Mol. Genet. 5:675-684(1996).
CC -1- FUNCTION: This protein shows both activity toward tyrosine-protein
CC phosphate as well as with serine/threonine-protein phosphate (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- TISSUE SPECIFICITY: Expressed predominantly in brain and lung.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X95518; CAA64772.1; -.
CC DR HSSP; Q16828; IMKP.
CC DR SMART; SM00195; DSC; 1.
CC MGD; MGI:106626; Dusp8.

```

```

DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR008343; MAPK_phosph.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANSE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolyase; Nuclear protein.
KM DOMAIN 23 138 RHODANSE.
FT DOMAIN 162 432 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 452 459 POLY-ARG.
FT DOMAIN 555 558 POLY-SER.
FT DOMAIN 559 576 POLY-GLY.
FT DOMAIN 577 600 POLY-SER.
FT DOMAIN 311 552 PRO-RICH.
FT ACT_SITE 246 246 PHOSPHOCYSTEINE INTERMEDIATE (BY
SQ SEQUENCE 663 AA; 68847 MW; 416F429A12C1FA7C CRC64;

Query Match 3.0%; Score 20; DB 1; Length 663;
Best Local Similarity 100.0%; Pred. No. 6.7e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 VHCLAGISRSATIAIAYIMK 261
Db 244 VHCLAGISRSATIAIAYIMK 263

RESULT 4
VHP1_CAEEL STANDARD; PRT; 619 AA.
ID VHP1_CAEEL 010038;
AC 010038;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein-tyrosine phosphatase vhp-1 (EC 3.1.3.48).
DE VHP-1 OR P08B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peldoderinae; Caenorhabditis.
ON NCBI_TaxID=6239;
RX MEDLINE=6239;
RA Chisoe S.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Non-receptor class dual specificity subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U23178; AAC46719.1; -.
CC DR PIR; T15969; T15969.
CC DR HSSP; Q16828; IMKP.
CC DR WormPep; P08B1.1; CE01899.
CC DR InterPro; IPR000340; DS_phosphatase.
CC DR InterPro; IPR000387; TYR_phosphatase.
CC DR Pfam; PF00782; DSPC; 1.
CC DR SMART; SM00195; DSC; 1.
CC DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.

```

DR PROSITE; PSS0056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolase.
 FT DOMAIN 200 269 PROTEIN-TYROSINE PHOSPHATASE.
 FT DOMAIN 92 95 POLY-SER.
 FT DOMAIN 351 354 POLY-SER.
 FT DOMAIN 465 472 POLY-SER.
 FT DOMAIN 483 488 POLY-SER.
 FT ACT_SITE 224 224 PHOSPHOCYSTEINE INTERMEDIATE (BY SIMILARITY)
 SQ SEQUENCE 619 AA; 66354 MW; 369B326F615D0529 CRC64;
 Query Match 1.4%; Score 9; DB 1; Length 619;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 243 HCLAGISRS 251
 Db 223 HCLAGISRS 231

RESULT 5
 ATP_PTRES STANDARD; PRT; 473 AA.
 AC 003080;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE ATP synthase beta chain (EC 3.6.3.14) (Fragment).
 GN ATPB.
 OS Pteridium esculentum.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Moniliformopses; Filicophyta; Filicopsida; Filicales;
 OC Dennstaedtiaceae; Pteridium.
 OX NCBI_TaxId=32102;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA WOLF P.G.;
 RT "Evaluation of atpB nucleotide sequences for phylogenetic studies of ferns and other pteridophytes";
 RL Am. J. Bot. 84:1429-1440(1997).
 CC -1- FUNCTION: Produces ATP from ADP in the presence of a proton gradient across the membrane. The beta chain is the catalytic subunit.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate + H(+) (out).
 CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic core - and CF(0) - the membrane proton channel. CF(1) has five subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0) has three main subunits: a, b and c.
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
 CC -1- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U93834; AAB51742.1; -
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR005722; ATP_synthet1_beta.
 DR InterPro; IPR000793; ATPase_a/b_C.
 DR InterPro; IPR000194; ATPase_a/b_centre.
 DR InterPro; IPR004100; ATPase_a/bN.
 DR InterPro; IPR009005; F1_ATPase_a/bN.
 DR Pfam; PF00006; ATP-synt_ab; 1.
 DR Pfam; PF00306; ATP-synt_ab_C; 1.
 DR Pfam; PF02874; ATP-synt_ab_N; 1.
 DR SMART; SM00382; AAA; 1.

DR TIGRFAM; TIGR01039; atpD; 1.
 DR PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
 KW ATP synthase; Chloroplast; Thylakoid; Membrane; CF(1);
 KW Hydrolase; Atp-binding; Hydrogen ion transport.
 FT NON_TER 1 1
 FT NP_BIND 172 179 ATP (POTENTIAL).
 FT NON_TER 473 473
 SQ SEQUENCE 473 AA; 50434 MW; 5F30596B8BE90029 CRC64;
 Query Match 1.1%; Score 7; DB 1; Length 473;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 63 VLITELI 69
 Db 180 VLITELI 186

RESULT 6
 LXC2_PROHE STANDARD; PRT; 478 AA.
 AC P29236;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Acyl-CoA reductase (EC 1.2.1.50).
 GN LXC.
 OS Photobacterium leiognathi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Photobacterium.
 OX NCBI_TaxId=658;
 RN [1]
 RA SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25521;
 RX MEDLINE=92007870; PubMed=1915359;
 RA Lee C.Y., Szilther R.B., Meighen E.A.;
 RT "The lux genes of the luminous bacterial symbiont, Photobacterium leiognathi, of the ponyfish. Nucleotide sequence, difference in gene organization, and high expression in mutant Escherichia coli.";
 RL Eur. J. Biochem. 201:161-167(1991).
 CC -1- FUNCTION: LXC IS THE FATTY ACID REDUCTASE ENZYME RESPONSIBLE FOR SYNTHESIS OF THE ALDEHYDE SUBSTRATE FOR THE LUMINESCENT REACTION CATALYZED BY LUCIFERASE.
 CC -1- CATALYTIC ACTIVITY: A long-chain aldehyde + CoA + NADP(+) = a long-chain acyl-CoA + NADPH.
 CC -1- PATHWAY: Bioluminescent fatty acid reduction system; third step.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M63594; AAA25616.1; -
 DR PIR; S17836; S17836.
 DR InterPro; IPR008670; LXC.
 DR Pfam; PF05893; LXC; 1.
 DR Luminase; Oxioreductase; NADP.
 SQ SEQUENCE 478 AA; 53713 MW; 4716FF699BF7FE4A CRC64;

Query Match 1.1%; Score 7; DB 1; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 117 SVHLLAG 123
 Db 145 SVHLLAG 151

RESULT 7
 MS14_ARATH

MS14 ARATH STANDARD; PRT: 507 AA.
 AC 022607; 042322; 042323; 093WF7; Q99LND;
 DT 15-UTL-1998 (Rel. 36, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE WD-40 repeat protein MS14.
 GN MS14 OR AT2G19520 OR F3P11.12.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopses.
 NC NCB1_TaxID=3702;
 RX SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RC MEDLINE=20083487; PubMed=10611797;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
 RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
 RA Goodman H.M., Somerville C.R., Coppenhaver G.P., Preuss D.,
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:761-768(1999).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Soultwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Nemman G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brocks S.Y., Carrinchi P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamliya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Thollogis A., Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 genome.";
 RL Science 302:842-846(2003).
 RN (3)
 RP SEQUENCE OF 4-507 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=99087420; PubMed=9872415;
 RA Kenzior A.L., Folk W.R.;
 RT "AtMS14 and RbAp48 WD-40 repeat proteins bind metal ions.";
 RL FEBS Lett. 440:425-429(1998).
 RN (4)
 RP SEQUENCE OF 1-142 AND 452-507 FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Green silique;
 RA Raynal M., Grellet F., Landie M., Meyer Y., Cooke R., Delenly M.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: Belongs to the WD-repeat RBA46/RBA48/MS11 family.
 CC -1- CAUTION: Ref.4 (CAA85542) sequence differs from that shown due to
 a frameshift in position 136.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AC005917; AAD10151.2; ALT_INIT.

DR EMBL; AY059799; AAL24281.1; -;
 DR EMBL; AY057655; AAL15286.1; -;
 DR EMBL; AY081447; AAM10009.1; -;
 DR EMBL; AF028711; AAD03340.1; -;
 DR EMBL; Z37286; CAA85542.1; ALT_FRAME.
 DR EMBL; Z37287; CAA85543.1; -;
 DR PIR; G84577; G84577.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 5.
 DR PRINTS; PR00320; GPROTEINRPT.
 DR ProDom; PD000018; WD40; 1.
 DR SMART; SM00320; WD40; 6.
 DR PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
 DR PROSITE; PS00882; WD_REPEATS_2; 3.
 DR PROSITE; PS02994; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat; Nuclear protein.
 FT REPEAT 162 193 WD 1.
 FT REPEAT 217 248 WD 2.
 FT REPEAT 290 321 WD 3.
 FT REPEAT 335 366 WD 4.
 FT REPEAT 384 415 WD 5.
 FT REPEAT 439 477 WD 6.
 FT DOMAIN 468 471 POLY-GLY.
 FT CONFLICT 89 89 W -> L (IN REF. 4; CAA85542).
 FT CONFLICT 126 126 V -> F (IN REF. 4; CAA85542).
 FT CONFLICT 202 202 A -> P (IN REF. 3).
 FT CONFLICT 270 270 T -> P (IN REF. 3).
 FT CONFLICT 463 463 D -> A (IN REF. 4; CAA85543).
 FT CONFLICT 489 489 V -> F (IN REF. 4; CAA85543).
 FT CONFLICT 494 494 E -> A (IN REF. 3).
 SQ SEQUENCE 507 AA; 55759 MW; C37F8000F8B33397 CR664;
 Query Match 1.1%; Score 7; DB 1; Length 507;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 302 TGAAGPK 308
 DB 20 TGAAGPK 26
 RESULT 8
 ID PUR1_CHICK STANDARD; PRT: 510 AA.
 AC P28173;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Amidophosphoribosyltransferase precursor (EC 2.4.2.14) (Glutamine
 DE phosphoribosylpyrophosphate amidotransferase) (ATPase) (GPAT).
 GN GPAT OR GPAT.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 NC NCB1_TaxID=9031;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC MEDLINE=91065928; PubMed=2123487;
 RA Zhou G., Dixon J.E., Zalkin H.;
 RT "Cloning and expression of avian glutamine
 RT phosphoribosylpyrophosphate amidotransferase. Conservation of a
 RT bacterial propeptide sequence supports a role for posttranslational
 RT processing.";
 RL J. Biol. Chem. 265:21152-21159(1990).
 RN (2)
 RP SEQUENCE OF 1-42 FROM N.A.
 RC MEDLINE=93330273; PubMed=8336716;
 RA Gavellas A., Dixon J.E., Brayton K.A., Zalkin H.;
 RT "Coexpression of two closely linked avian genes for purine nucleotide
 RT synthesis from a bidirectional promoter.";
 RL Mol. Cell. Biol. 13:4784-4792(1993).
 RN (3)

```

RP ERRATUM.
RA Gavalas A., Dixon J.E., Brayton K.A., Zalkin H.;
RL Mol. Cell. Biol. 13:7977-7977(1993).
CC
CC -1- CATALYTIC ACTIVITY: 5-phospho-beta-D-riboylamine + diphosphate +
CC L-glutamate = L-glutamine + 5-phospho-alpha-D-ribose 1-diphosphate
CC + H(2)O.
CC
CC -1- COFACTOR: Binds 1 magnesium ion and 1 4Fe-4S cluster per subunit
CC (by similarity).
CC
CC -1- PATHWAY: De novo purine biosynthesis; first step.
CC
CC -1- SIMILARITY: In the C-terminal section, belongs to the
CC purine/pyrimidine phosphoribosyltransferase family.
CC
CC -1- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: M60069; AAA62736.1; -.
DR EMBL: LI2533; AAA17895.1; -.
DR PIR: A38337; A38337.
DR HSP: P00497; 1A00.
DR MEROPS: C44.001; -.
DR InterPro: IPR005854; Amd_phospho_trans.
DR InterPro: IPR000583; GATase_2.
DR InterPro: IPR002375; Pr_Py_TP_transf.
DR InterPro: IPR000836; PRTtransferase.
DR Pfam: PF00310; GATase_2; 1.
DR Pfam: PF00156; Pribosyltran; 1.
DR TIGRFAMs: TIGR01134; purF; 1.
DR PROSITE: PS00103; PUR_PYR_PR_TRANSFER; 1.
DR PROSITE: PS00443; GATASE_TYPE_II; 1.
DR Purine biosynthesis; transferase; Glycosyltransferase;
DR Glutamine amidotransferase; Allosteric enzyme; Metal-binding;
DR Magnesium; Iron-sulfur; 4Fe-4S.
DR PROPEP 1 11 PROBABLE.
DR CHAIN 12 510 AMIDOPHOSPHORIBOSYLTRANSFERASE.
DR ACT SITE 12 12 GATASE (BY SIMILARITY).
DR METAL 280 280 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
DR METAL 327 327 MAGNESIUM (BY SIMILARITY).
DR METAL 389 389 MAGNESIUM (BY SIMILARITY).
DR METAL 390 390 MAGNESIUM (BY SIMILARITY).
DR METAL 426 426 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
DR METAL 496 496 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
DR METAL 499 499 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
DR SEQUENCE 510 AA; 56257 MW; F4371FE1FEC7C744 CRC64;

Query Match 1.1%; Score 7; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 594 VYSVRR 600
DB 299 VYSVRR 305

RESULT 9
HEMK_RICCN STANDARD; PRT; 524 AA.
ID HEMK_RICCN
AC Q92613;
DT 10-OCT-2003 (Rel. 42; Created)
DT 10-OCT-2003 (Rel. 42; Last sequence update)
DT 10-OCT-2003 (Rel. 42; Last annotation update)
DE Bifunctional methyltransferase [includes: Hemk protein homolog
DE (EC 2.1.1.-) (M.RcoHemK) / tRNA [guanine-N(7)-methyltransferase
DE (EC 2.1.1.33) (tRNA (m7G46)-methyltransferase)].
OS HEMK OR RC1314.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;

```

```

OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audifren P., Fournier P.-E., Barbe V.,
RA Samon D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
CC
CC -1- FUNCTION: Catalyzes the formation of N(7)-methylguanine at
CC position 46 (m7G46) in tRNA (By similarity).
CC
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
CC homocysteine + tRNA containing N(7)-methylguanine.
CC
CC -1- SIMILARITY: In the N-terminal section, belongs to the hemk family.
CC
CC -1- SIMILARITY: In the C-terminal section, belongs to the
CC methyltransferase superfamily. Trmb family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: AE008678; AAL03852.1; -.
DR HAMAP: MF_01057; fused; 1.
DR InterPro: IPR004395; Cons_hypoth91.
DR InterPro: IPR004556; HemK.
DR InterPro: IPR003358; Methyltransf_4.
DR InterPro: IPR002052; NG_Mcase.
DR InterPro: IPR000051; SAM_bind.
DR Pfam: PF02390; Methyltransf_4; 1.
DR TIGRFAMs: TIGR00536; hemk fam; 1.
DR TIGRFAMs: TIGR00091; TIGR00091; 1.
DR PROSITE: PS00092; NG_MCASE; UNKNOWN 1.
DR Transferrase; Methyltransferase; tRNA processing; Complete proteome.
DR DOMAIN 1 306 HEMK.
DR DOMAIN 307 524 TRNA (GUANINE-N(7)-METHYLTRANSFERASE.
FT SEQUENCE 524 AA; 60079 MW; 3E574DC6F7DFDC28 CRC64;

Query Match 1.1%; Score 7; DB 1; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 ELIQNG 184
DB 473 ELIQNG 479

RESULT 10
TCEI_AVESA STANDARD; PRT; 535 AA.
ID TCEI_AVESA
AC P40412;
DT 01-FEB-1995 (Rel. 31; Created)
DT 01-FEB-1995 (Rel. 31; Last sequence update)
DT 01-OCT-1996 (Rel. 34; Last annotation update)
DE T-complex protein 1, epsilon subunit (TCP-1-epsilon) (CCT-epsilon)
DE (TCP-X19).
OS Avena sativa (Oat).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Aveneae; Avena.
OX NCBI_TaxID=4498;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Dewi; TISSUE=coleoptile, and Mesocotyl;
RX MEDLINE=94085629; PubMed=7903257;
RA Elmann B., Krenz W., Mummert E., Schaefer E.;
RA "Two Tcp-1-related but highly divergent gene families exist in oat
RT encoding proteins of assumed chaperone function.";

```

```
RL FEBS Lett. 336:313-316(1993).
CC -1- FUNCTION: Molecular chaperone; assist the folding of proteins upon
CC ATP hydrolysis. Known to play a role, in vitro, in the folding of
CC actin and tubulin.
CC -1- SUBUNIT: Hetero-oligomeric complex of about 850 to 900 kDa that
CC forms two stacked rings, 12 to 16 nm in diameter.
CC -1- SIMILARITY: Belongs to the TCP-1 chaperonin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75777; CAA53396.1; -.
CC PIR; S40461; S40461.
CC HSSP; P48425; 1A6D.
CC InterPro; IPR002194; Chaperonin TCP-1.
CC InterPro; IPR001844; Chaperonin Cpn60.
CC InterPro; IPR002423; Cpn60/TCP-1.
CC Pfam; PF00118; Cpn60 TCP1_1.
CC PRINTS; PR00298; CHAPERONIN60.
CC PROSITE; PS00750; TCP1_1; 1.
CC PROSITE; PS00751; TCP1_2; 1.
CC PROSITE; PS00995; TCP1_3; 1.
CC Chaperone; ATP-binding; Multigene family.
SQ SEQUENCE 535 AA; 58899 MW; 6A8F847CA891BC32 CRC64;

Query Match 1.1%; Score 7; DB 1; Length 535;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 KHKVDID 80
DB 259 KHKVDID 265

RESULT 11
TCE2_AVEA
ID TCE2_AVEA STANDARD; PRT; 535 AA.
AC PS411;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE T-complex protein 1, epsilon subunit (TCP-1-epsilon) (CCT-epsilon)
DE (TCP-K36).
OS Avena sativa (oat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
OC Avenaeae; Avena.
OX NCBI_TaxID=498;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Pearl; TISSUE=coleoptile, and Mesocotyl;
RA MEDLINE=94085629; PubMed=7903257;
RA Ehlmann B., Krenz M., Mumme E., Schaefer E.;
RT "Two Tcp-1-related but highly divergent gene families exist in oat
RT encoding proteins of assumed chaperone function.";
RL FEBS Lett. 336:313-316(1993).
CC -1- FUNCTION: Molecular chaperone; assist the folding of proteins upon
CC ATP hydrolysis. Known to play a role, in vitro, in the folding of
CC actin and tubulin.
CC -1- SUBUNIT: Hetero-oligomeric complex of about 850 to 900 kDa that
CC forms two stacked rings, 12 to 16 nm in diameter.
CC -1- SIMILARITY: Belongs to the TCP-1 chaperonin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75778; CAA53397.1; -.
CC PIR; S40462; S40462.
CC HSSP; P48424; 1A6D.
CC InterPro; IPR002194; Chaperonin TCP-1.
CC InterPro; IPR001844; Chaperonin Cpn60.
CC InterPro; IPR002423; Cpn60/TCP-1.
CC Pfam; PF00118; Cpn60 TCP1_1.
CC PRINTS; PR00298; CHAPERONIN60.
CC PROSITE; PS00750; TCP1_1; 1.
CC PROSITE; PS00751; TCP1_2; 1.
CC PROSITE; PS00995; TCP1_3; 1.
CC Chaperone; ATP-binding; Multigene family.
SQ SEQUENCE 535 AA; 59034 MW; 02761226F5F17B81 CRC64;

Query Match 1.1%; Score 7; DB 1; Length 535;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 KHKVDID 80
DB 259 KHKVDID 265

RESULT 12
TCE2_ARATH
ID TCE2_ARATH STANDARD; PRT; 535 AA.
AC 004450;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE T-complex protein 1, epsilon subunit (TCP-1-epsilon) (CCT-epsilon).
DE A11G24510 OR F21J9.12.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=21016719; PubMed=1130712;
RA Theologis A., Becker J.R., Palm C.D., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin B.,
RA Kim C.U., Koo H.L., Kremetskaia I., Kutz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltsecher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
RA Uteback T., Van Aken S., Vayenberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana ";
RL Nature 408:816-820(2000).
CC -1- FUNCTION: Molecular chaperone; assist the folding of proteins upon
CC ATP hydrolysis. Known to play a role, in vitro, in the folding of
CC actin and tubulin.
CC -1- SUBUNIT: Hetero-oligomeric complex of about 850 to 900 kDa that
```

CC		forms two stacked rings, 12 to 16 nm in diameter.
CC	-I-	SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC	-I-	SIMILARITY: Belongs to the TCP-1 chaperonin family.
CC		-----
CC		This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC		the European Bioinformatics Institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC		or send an email to license@isb-sib.ch).
CC		-----
DR	EMBL:	AC000103; AAF97977.1; .
DR	HSSP:	P48424; IABD.
DR	InterPro:	IPIR002194; Chaperonin_TCP-1.
DR	InterPro:	IPIR001844; Chaprinin_Cpm60.
DR	InterPro:	IPIR002423; Cpn60/TCP-1.
DR	InterPro:	IPIR008950; GROEL-ATFase.
DR	Pfam:	PF00118; cpn60_TCP1_1.
DR	PRINTS:	PR00298; CHAPERONIN60.
DR	PRINTS:	PR00304; TCOMPLEXTCP1.
DR	PROSITE:	PS00750; TCP1_1; 1.
DR	PROSITE:	PS00751; TCP1_2; 1.
DR	PROSITE:	PS00995; TCP1_3; 1.
KW	Chaperone;	ATP-binding; Multigene family.
SO	SEQUENCE	535 AA; 59384 MW; 1CB56343A3AFIDC3 CRC64;
Oy		
Dz		
Oy	74 KHKVDID 80	1.1%; Score 7; DB 1; Length 535; Best Local Similarity 100.0%; Pred.No.1.le+02; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	259 KHKVID 265	
RESULT 13		
ID	NOXI_HUMAN	STANDARD; PRT; 564 AA.
AC	Q9Y5S8; O95691;	
DT	30-MAY-2000 (Rel. 39, Created)	
DT	30-MAY-2000 (Rel. 39, Last sequence update)	
DT	15-MAR-2004 (Rel. 43, Last annotation update)	
DE	NADPH oxidase homolog 1 (NOX-1) (NMDH/NADPH mitogenic	
DE	oxidase subunit P65-MOX) (Mitogenic oxidase 1) (MOX1).	
GN	NOX1 OR NOX1.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A. (ISOFORM NOH-1L).	
RC	TISSUE=Colon epithelium;	
RX	MEDLINE=99413719; PubMed=10485709;	
RA	Sub Y.-A., Arnold R.S., Lassegue B., Shi J., Xu X., Sorensen D.,	
RA	Chung A.B., Griending K.K., Lambeth J.D.;	
RL	"Cell transformation by the superoxide-generating oxidase Moxi.";	
RL	Nature 401:79-82(1999).	
RN	[2]	
RP	SEQUENCE FROM N.A. (ISOFORMS NOH-1L, NOH-1LV AND NOH-1S).	
RX	MEDLINE=20082959; PubMed=10615049;	
RA	Banfi B., Maturana A., Jaconí S., Arnaudau S., Laforge T., Sinha B.,	
RA	Liget B., Demareux N., Krause K.-H.;	
RT	"A mammalian H+ channel, generated through alternative splicing of the	
RT	NADPH oxidase homolog NOH-1.";	
RL	Science 287:138-142(2000).	
RN	[3]	
RP	SEQUENCE OF 16-564 FROM N.A. (ISOFORM NOH-1L).	
RL	Lloyd D.;	
CC	Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.	
CC	-I- FUNCTION: NOH-1S is a voltage-gated proton channel that mediates	
CC	the H(+) currents of resting phagocytes and other tissues. It	
CC	participates in the regulation of cellular pH and is blocked by	

	CC	zinc_NOH-1L is a pyridine nucleotide-dependent oxidoreductase
	CC	that generates superoxide and might conduct H(+) ions as part of
	CC	its electron transport mechanism; whereas NOH-1S does not contain
	CC	an electron transport chain.
	-I-	COPACOFER: NADP and FAD (Potential).
	CC	-I- SUBCELLULAR LOCATION: Integral membrane protein (potential).
	-I-	ALTERNATIVE PRODUCTS:
	CC	Event-Alternative splicing; Named isoforms=3;
	CC	Name=NOH-1L;
	CC	IsoId=Q9Y5S8-1; Sequence=Displayed;
	CC	Name=NOH-1S;
	CC	IsoId=Q9Y5S8-2; Sequence=VSP_001577; VSP_001578;
	CC	Name=NOH-1LV;
	CC	IsoId=Q9Y5S8-3; Sequence=VSP_001579;
	-I-	TISSUE SPECIFICITY: NOH-1L is detected in colon, uterus, prostate,
	CC	and colon carcinoma, but not in peripheral blood leukocytes. NOH-
	CC	-1S is detected only in colon and colon carcinoma cells.
	-I-	SIMILARITY: Belongs to the FRE / CYBB family.
	CC	-----
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
	CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
	CC	the European Bioinformatics Institute. There are no restrictions on its
	CC	use by non-profit institutions as long as its content is in no way
	CC	modified and this statement is not removed. Usage by end users for commercial
	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licenses@isb.sib.ch). -----
	DR	EMBL AF12763; AAB38133.1; .
	DR	EMBL AF16632C; AAFF3232.1; .
	DR	EMBL AF16632F; AAF32323.1; .
	DR	EMBL AF16632B; AAF32324.1; .
	DR	EMBL Z83819; CABO6073.1; ALT_SEQ.
	DR	GeneW; HGNC:7889; NOX1.
	DR	MIM: 300225; .
	DR	GO GO:0016021; C:integral to membrane; NAS.
	DR	GO GO:0016175; P:superoxide-generating NADPH oxidase activity; TAS.
	DR	GO GO:0030171; F:voltage-gated proton channel activity; TAS.
	DR	GO GO:0006118; P:electron transport; NAS.
	DR	GO GO:0006746; P:FADH2 metabolism; NAS.
	DR	GO GO:0006739; P:NADPH metabolism; NAS.
	DR	GO GO:0015992; P:proton transport; TAS.
	DR	InterPro; IPR002916; Ferric reduct.
	DR	InterPro; IPR000778; GP1Phox.
	DR	Pfam; PF01794; Ferric_reduct; 1.
	KM	Oxidoreductases; NADP; Electron transport; Transmembrane; FAD; Heme; Glycoprotein; Voltage_gated channel; Ionic channel;
	KW	Alternatieve aplicing.
	FT	DNAIN 1 9
	FT	TRANSEM 10 30
	FT	DNAIN 31 44
	FT	TRANSEM 45 72
	FT	DNAIN 73 102
	FT	TRANSEM 103 123
	FT	DNAIN 124 168
	FT	TRANSEM 169 189
	FT	DNAIN 190 206
	FT	TRANSEM 207 227
	FT	DNAIN 228 396
	FT	TRANSEM 397 417
	FT	DNAIN 418 564
	FT	NP BIND 338 344
	FT	METAL 101 101
	FT	METAL 115 115
	FT	METAL 209 209
	FT	METAL 221 221
	FT	CARBOND 162 162
	FT	CARBOND 236 236
	FT	VASPLIC 159 190
	VARSPIC 191 564	
	Missing (in isoform NOH-1S).	/Fld=vsp_001577. Missing (in isoform NOH-1S).

FT VARSPLIC 433 481 /FTId=VSP 001578.
 FT Missing (In isoform NOH-1LV).
 FT CONFLICT 173 173 I -> V (IN REF. 2).
 SQ SEQUENCE 564 AA, 64870 MW, C3BE290F456DBC9A CRC64;
 Query Match 1.1%; Score 7; DB 1; Length 564;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 644 BELRGV 650
 DB 460 BELRGV 466

RESULT 14
 SCC1 YEAST STANDARD; PRT; 566 AA.
 AC 012158; 005325;
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Sister chromatid cohesion protein 1.
 GN MCD1 OR SCC1 OR RHC21 OR PDS3 OR YDL003W OR YD8119.04.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
 RC STRAIN=W303;
 RX MEDLINE=97474309; PubMed=9335333;
 RA Michaelis C., Clook R., Naemlyth K.;
 RT "Cohesins: chromosomal proteins that prevent premature separation of
 RL sister chromatids." Cell 91:35-45(1997).
 RN [2]
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
 RC STRAIN=S288c / FY1678;
 RX MEDLINE=97474310; PubMed=9335334;
 RA Guacci V., Koshland D., Strunnikov A.V.;
 RT "A direct link between sister chromatid cohesion and chromosome
 RL condensation revealed through the analysis of MCD1 in S. cerevisiae." Cell 91:47-57(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / FY1678;
 RX MEDLINE=97113263; PubMed=9169867;
 RA Jacq C., Alt-Moerbe J., Andre B., Arnold W., Bahr A., Ballestra J.P.G.,
 RA Barques M., Baron L., Becker A., Biteau N., Bloescher H., Blugnon C.,
 RA Bosgrove T., Brandt P., Bruckner M., Bultrago M.J., Coster F.,
 RA Delvaux T., del Rey F., Dujon B., Elde L.G., Garcia-Cantalejo J.M.,
 RA Goffeau A., Gomez-Perez A., Granotier C., Hanemann V., Harkett T.,
 RA Hochstedt J.D., Jaeger W., Jimenez A., Joniaux J.-L., Kraemer C.,
 RA Kuester H., Laamanen P., Legros Y., Louis E.J., Moeller-Rieker S.,
 RA Monnet A., Moro M., Mueller-Auer S., Nusbaumer B., Particio N.,
 RA Paulin L., Perez-Alonso M., Perez-Ortin J.E., Pohl T.M.,
 RA Prydz H., Purnelle B., Rasmussen S.W., Remacha M., Revuelta J.L.,
 RA Rieger M., Salom D., Saluz H.P., Saiz J.E., Saren A.-M., Schaefer M.,
 RA Scherfe M., Schmidt E.R., Schneider C., Scholler P., Schwarz S.,
 RA Utratazu L.A., Verhasselt P., Vissers S., Voet M., Voickaert G.,
 RA Wagner G., Wandut R., Wedler E., Wedler H., Woelfl S., Harris D.E.,
 RA Bowman S., Brown D., Churcher C.M., Connor R., Dedman K., Gentles S.,
 RA Hamlin N., Hunt S., Jones L., McDonald S., Murphy L., Niblett D.,
 RA Odell C., Oliver K., Rajandream M.A., Richards C., Shore L.,
 RA Walsh S.V., Barrell B.G., Dietrich F.S., Mulligan J.T., Allen E.,
 RA Araujo R., Aviles E., Berno A., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Hunnicke-Smith S., Hyman R.W., Komp C.,
 RA Lashkari D., Lew H., Lin D., Mosedale D., Nakahara K., Namath A.,
 RA Oether P., Oh C., Petel F.X., Roberts D., Schramm S., Schroeder M.,
 RA Schjoren T., Shroff N., Wiant A., Yelton M.A., Botstein D.,
 RA Davis R.W., Johnston M., Andrews S., Brinkman R., Cooper J., Ding H.,
 Du Z., Favell A., Fulton L., Gattung S., Greco T., Hallsworth K.,

RA Hawkins J., Hillier L.W., Jier M., Johnson D., Johnston L.,
 RA Kirsten J., Kucaba T., Langston V., Latreille P., Le T., Mardis E.,
 RA Meneses S., Miller N., Nham M., Pauley A., Peluso D., Rickin L.,
 RA Riles L., Tach A., Treaskis E., Vignati D., Wilcox L., Woldman P.,
 RA Vaudin M., Waisson R., Waterston R., Albertmann K., Hant J., Heumann K.,
 RA Kleine K., Mewes H.-W., Zollner A., Zaccaria P.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IV." J
 RL Nature 387:75-78(1997).
 RN [4]
 RP FUNCTION.
 RX MEDLINE=98151773; PubMed=9491073;
 RA Heo S.-J., Tatebayashi K., Kato J., Ikeda H.;
 RT "The RHC21 gene of budding yeast, a homologue of the fission yeast
 RT rad21+ gene, is essential for chromosome segregation." J
 RL Mol. Genet. 257:149-156(1998).
 RN [5]
 RP SUBCELLULAR LOCATION, INTERACTION WITH IRR1, IDENTIFICATION IN A
 RP COHESIN COMPLEX WITH SMC1, SMC3 AND IRR1, AND INTERACTION OF THE
 RP COHESIN COMPLEX WITH SCC2.
 RX MEDLINE=99145468; PubMed=9990856;
 RA Toch A., Clook R., Uhlmann F., Galova M., Schleiffer A., Naemlyth K.;
 RT "Yeast cohesin complex requires a conserved protein, Escp1(Ctf17), to
 RT establish cohesion between sister chromatids during DNA replication." J
 RL Genes Dev. 13:320-333(1999).
 RN [6]
 RP CLEAVAGE BY ESPI, FUNCTION, AND MUTAGENESIS OF ARG-180 AND ARG-268.
 RX MEDLINE=99330041; PubMed=10403247;
 RA Uhlmann F., Lottspeich F., Naemlyth K.;
 RT "Sister-chromatid separation at anaphase onset is promoted by cleavage
 RT of the cohesin subunit Scc1." Nature 400:37-42(1999).
 RN [7]
 RP PHOSPHORYLATION BY CDC5, AND MUTAGENESIS OF SER-175 AND SER-263.
 RX MEDLINE=21264235; PubMed=11371343;
 RA Alexandru G., Uhlmann F., Mechtler K., Poupard M.-A., Naemlyth K.;
 RT "Phosphorylation of the cohesin subunit Scc1 by Polo/Cdc5 kinase
 RT regulates sister chromatid separation in yeast." Cell 105:459-472(2001).
 RL Cell 105:459-472(2001).
 RN [8]
 RP IDENTIFICATION IN A COHESIN COMPLEX WITH SMC1, SMC3 AND IRR1, AND
 RP STRUCTURE.
 RX MEDLINE=21980168; PubMed=11983169;
 RA Haering C.H., Loewe J., Hochwagen A., Naemlyth K.;
 RT "Molecular architecture of SMC proteins and the yeast cohesin
 RT complex." Mol. Cell 9:773-784(2002).
 RL Mol. Cell 9:773-784(2002).
 CC -1- FUNCTION: Cleavable component of the cohesin complex involved in
 CC chromosome cohesion during cell cycle. The cohesin complex is
 CC required for the cohesion of sister chromatids after DNA
 CC replication. The cohesin complex apparently forms a large
 CC proteinaceous ring within which sister chromatids can be trapped.
 CC At metaphase-anaphase transition, this protein is cleaved by ESPI
 CC and dissociates from chromatin, allowing sister chromatids to
 CC segregate.
 CC -1- SUBUNIT: Interacts directly with IRR1/SCC3 in cohesin complex.
 CC Cohesin complexes are composed of the SMC1 and SMC3 heterodimer
 CC attached via their hinge domain, MCD1/SCC1 which link them, and
 CC IRR1, which interacts with MCD1. The cohesin complex also
 CC interacts with SCC2, which is required for its association with
 CC chromosomes.
 CC -1- SUBCELLULAR LOCATION: Nuclear protein. Associates with chromatin.
 CC Before prophase it is scattered along chromosome arms. During
 CC prophase, most of cohesin complexes dissociate from chromatin
 CC except at centromeres, where cohesin complexes remain. At
 CC anaphase, it is cleaved by ESPI, leading to the dissociation of
 CC the complex from chromosomes, allowing chromosome separation.
 CC -1- DOMAIN: The C-terminal part associates with the head of SMC1,
 CC while the N-terminal part binds to the head of SMC3.
 CC -1- PTM: Cleaved by ESPI at the onset of anaphase.
 CC -1- PTM: Phosphorylation by CDC5/Polo-like kinase at the onset of
 CC anaphase. Phosphorylation takes places at proximity to cleavage
 CC sites and is required for an efficient cleavage by ESPI.
 CC -1- SIMILARITY: Belongs to the SCC1/RAD21 family.

```

CC ----- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL; Y14280; CAA74657.1; -
DR EMBL; U23759; AAB38803.1; -
DR EMBL; Z48008; CAA68058.1; -
DR EMBL; Z48432; CAA68356.1; -
DR EMBL; Z74051; CAA68559.1; -
DR PIR; S50979; S50979; -
DR GeneOnline; 140245; -
DR SGD; S0002161; MCD1; -
DR GO; GO:0007076; P:mitotic chromosome condensation; IMP.
DR GO; GO:0007064; P:mitotic sister chromatid cohesion; IMP.
DR InterPro; IPR006909; Rad21_Rech.
DR InterPro; IPR006910; Rad21_Rech_N.
DR Pfam; PF04824; Rad21_Rech_1.
DR Pfam; PF04825; Rad21_Rech_N; 1.
DR Microsil; Cell cycle; Chromosome partition; Nuclear protein;
KW Phosphorylation.
FT DOMAIN 258 261 POLY-ASP.
FT SITE 180 181 CLEAVAGE (BY ESP1).
FT SITE 268 269 CLEAVAGE (BY ESP1).
FT MOD_RES 175 175 PHOSPHORYLATION (BY CDC5).
FT MOD_RES 263 263 PHOSPHORYLATION (BY CDC5).
FT MUTAGEN 175 175 S->A: REDUCES PHOSPHORYLATION. ABOLISHES
      PHOSPHORYLATION WHEN ASSOCIATED WITH A-
      263.
FT FT R->D: ABOLISHES CLEAVAGE BY ESP1 WHEN
      ASSOCIATED WITH D-268.
FT MUTAGEN 180 180 S->A: REDUCES PHOSPHORYLATION. ABOLISHES
      PHOSPHORYLATION WHEN ASSOCIATED WITH A-
      175.
FT FT R->D: ABOLISHES FIRST CLEAVAGE BY ESP1.
      ABOLISHES ALL CLEAVAGE BY ESP1 WHEN
      ASSOCIATED WITH D-180.
FT FT
FT FT
SQ SEQUENCE 566 AA; 63290 MW; EB6C7CA33BAC208F CRC64;

Query Match          1.1%; Score 7; DB 1; Length 566;
Best local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 633 GESIMSE 639
   |||||
Db 271 GESIMSE 277

RESULT 15
MALT_AEDAE
AD MALT_AEDAE STANDARD; PRT; 579 AA.
AC PI3080;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable maltase precursor (EC 3.2.1.20).
GN MALI OR MAL I.
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
OX NCBI_TaxID=71557;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=Rockefeller, and Bahama; TISSUE=Salivary gland;
RX MEDLINE=89252923; Pubmed=2470653;
RA James A.A., Blackmer K., Ractopai J.V.;
RT "A salivary gland-specific, maltase-like gene of the vector mosquito,
RL Aedes aegypti."
Gene 75:73-83(1989).
```

Qy	Db	577 ASYSAYS 583	414 ASYSAYS 420
1.1%	Score 7; DB 1; Length 579;		
Best Local Similarity	100.0%; Pred. No. 1.2e+02;		
Matches	7; Conservative	0; Mismatches	0; Indels
		0; Gaps	0;

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbhnowtsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
 RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambute R., Fumelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelure V., Mottier S.,
 RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucase M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado J., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armerstrong J., Forzbach S.L.,
 RA Cernetti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
 RA Shpikovski G.V., Useery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of *Schizosaccharomyces pombe*,"
 RL Nature 415:871-880(2002).
 RN [2]
 RP SEQUENCE OF 108-583 FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=95161095; PubMed=7857672;
 RA Ishiguro J., Uhara Y., Kawahara K.;
 RT "Molecular cloning and characterization of a fission yeast gene
 responsible for supersensitivity to the spindle poison, isopropyl
 N-3-chlorophenyl carbamate.";
 RL Jpn. J. Genet. 69:671-678(1994).
 RN [3]
 RP REVISIONS.
 RA Ishiguro J.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Responsible for supersensitivity to the spindle poison,
 isopropyl N-3-chlorophenyl carbamate.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: 299260; CAB16391.1; -;
 DR EMBL: AB017490; BAA33049.1; -;
 DR PIR: T11624; T11624.
 DR GeneDB: SPombe; SPAC3A11.02; -;
 DR InterPro: IPR000571; Znf_CCH.
 DR Pfam: PF00642; zf-CCH; 2.
 DR SMART: SM00356; Znf_C3H1; 2.
 DR SMART: 583 AA; 62807 MM; 905022C7C06E4271 CRC64;
 SQ SEQUENCE 583 AA; 62807 MM; 905022C7C06E4271 CRC64;
 QY 408 SYSASNA 414
 DB 298 SYSASNA 304
 RESULT 17
 ID ILV3_YEAST STANDARD; PRT; 585 AA.
 AC P39522;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Dihydroxy-acid dehydratase, mitochondrial precursor (EC 4.2.1.9) (DAD)
 GN (2,3-dihydroxy acid hydrolase).
 DN ILV3 OR YJR016C OR J1450.
 OS *Saccharomyces cerevisiae* (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC *Saccharomycetaceae*; *Saccharomycetaceae*; *Saccharomycetes*.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S28bc / F11679;
 RA de Haan M., Smits P.H.M., Grievell L.A.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 82-585 FROM N.A.
 RC STRAIN=X2180;
 RX MEDLINE=94131281; PubMed=8299945;
 RA Velasco J.A., Cansado J., Pena M.C., Kawakami T., Laborda J.,
 RA Notario V.;
 RT "Cloning of the dihydroxyacid dehydratase-encoding gene (ILV3) from
 RT *Saccharomyces cerevisiae*,"
 RL Gene 137:179-185(1993).
 CC -1- CATALYTIC ACTIVITY: 2,3-dihydroxy-3-methylbutanoate = 3-methyl-2-
 CC oxobutanoate + H(2)O.
 CC -1- COFACTOR: Binds 1 4Fe-4S cluster (Potential).
 CC -1- PATHWAY: Valine and isoleucine biosynthesis; fourth step.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SIMILARITY: Belongs to the ilvD / edd family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X87611; CAA60939.1; -;
 DR EMBL: 249516; CAA89540.1; -;
 DR EMBL: L13975; AAA34568.1; ALT_TERM.
 DR PIR: S55205; S55205.
 DR GeneOnline: 141853; -;
 DR SGD: S0003777; ILV3.
 DR InterPro: IPR004404; ILV3.
 DR InterPro: IPR000581; ILV3_EDD_family.
 DR Pfam: PF00920; ILV3_EDD_1.
 DR ProDom: PD002691; ILV3_EDD_family; 1.
 DR TIGRFAMs: TIGR00110; ILV3_1.
 DR PROSITE: PS00886; ILV3_EDD_1; 1.
 DR PROSITE: PS00887; ILV3_EDD_2; 1.
 DR TRANSIT peptide; Iron; Iron-sulfur; 4Fe-4S.
 KW Branched-chain amino acid biosynthesis; Lyase; Mitochondrion;
 KM Transit peptide; Iron; Iron-sulfur; 4Fe-4S.
 FT TRANSIT 1 585
 FT CHAIN ? ?
 FT METAL 143 143
 FT METAL 221 221
 FT METAL 82 89
 FT CONFLICT 238 238
 FT CONFLICT 242 242
 FT CONFLICT 492 492
 FT CONFLICT 520 520
 FT CONFLICT 551 551
 FT CONFLICT 551 551
 FT CONFLICT 551 551
 SQ SEQUENCE 585 AA; 62861 MM; 35AB3C679BA658D0 CRC64;
 QY 235 ASNGCVL 241
 DB 577 ASNGCVL 583
 RESULT 18
 ID ILV3_YEAST STANDARD; PRT; 591 AA.
 AC P14347;
 DT 01-JAN-1990 (Rel. 13, Created)

```

DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Protein BFRF2.
GN BFRF2.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OC NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=4270667; PubMed=6087149;
RA Beer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seglin C.,
RA Tufnell P.S., Barrett B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).
CC -1- SIMILARITY: Belongs to the herpesviruses UL49 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01555; CAA24880.1; ALT_INIT.
DR InterPro; IPR004339; UL49.
DR Pfam; PF03117; UL49; 1.
SQ SEQUENCE 591 AA; 63977 MW; EE63FF1E0721912E CRC64;

Query Match
Best Local Similarity 1.1%; Score 7; DB 1; Length 591;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 324 PAVSEGG 330
DB 313 PAVSEGG 319

RESULT 19
Y745_ARATH STANDARD; PRT; 597 AA.
AC P59278; Q9C812;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hypothetical protein At1g51745.
GN At1g51745 OR F19C24.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egtu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
RA Miltschev J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shlim P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,

```

```

RA Uterback T., Van Aken S., Vaysberg M., Vyotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 406:816-820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22954850; PubMed=14593172;
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Souvick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.W., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari P.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Khan S., Koeseema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tanase R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome.";
RL Science 302:842-846(2003).
CC -1- SIMILARITY: Contains 1 PWM domain.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AC025294; AAG50888.1; ALT_SEQ.
DR EMBL; BT002989; AAO2798.1; -.
DR InterPro; IPR000313; PWM domain.
DR PROSITE; PS50812; PWM; 1.
FM Hypothetical protein.
FT DOMAIN 16 78 PWM.
FT CONFLICT 268 268 M -> R (IN REF. 1).
SQ SEQUENCE 597 AA; 65044 MW; E59141F5B9DFD79F CRC64;

Query Match
Best Local Similarity 1.1%; Score 7; DB 1; Length 597;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 IEXAKS 236
DB 87 IEXAKS 93

RESULT 20
HNFA_MOUSE STANDARD; PRT; 628 AA.
AC P223E1;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hepatocyte nuclear factor 1-alpha (HNF-1A) (Liver-specific
DE transcription factor 1F-B1) (LFB1).
GN TCF1 OR HNFA OR HNF-1A OR HNF-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91088607; PubMed=2263635;
RA Kuo C.U., Conley P.B., Hsieh C.L., Francke U., Crabtree G.R.;
RT "Molecular cloning, functional expression, and chromosomal

```


RT localization of mouse hepatocyte nuclear factor 1,"
 CC Proc. Natl. Acad. Sci. U.S.A. 87:9838-9842(1990).
 CC -1- FUNCTION: Required for the expression of several liver specific
 CC genes. Binds to the inverted palindrome 5'-GTTAATATTAC-3'.
 CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Liver.
 CC -1- SIMILARITY: Belongs to the HNF1 homeobox family.
 CC -1- SIMILARITY: Contains 1 homeobox domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; M57966; AAA37821.1; -
 CC DR PIR; A39262; A39262.
 CC DR PDB; 1F93; 20-SEP-00.
 CC DR PDB; 1G2Y; 17-JAN-01.
 CC DR PDB; 1G2Z; 17-JAN-01.
 CC DR PDB; 1G39; 17-JAN-01.
 CC DR PDB; 1J86; 11-JUL-01.
 CC DR PDB; 1LFB; 31-OCT-93.
 CC DR TRASPAC; T01211; -
 CC DR MGD; MGI:98504; Tcf1.
 CC DR GO; GO:0005634; C:nucleus; IDA.
 CC DR InterPro; IPR006899; HNF-1_N.
 CC DR InterPro; IPR006898; HNF1A_C.
 CC DR InterPro; IPR006897; HNF1b_C.
 CC DR InterPro; IPR001356; Homeobox.
 CC DR Pfam; PF04814; HNF-1_N; 1.
 CC DR Pfam; PF04813; HNF-1A_C; 1.
 CC DR Pfam; PF04812; HNF-1B_C; 1.
 CC DR SMART; SMO0389; HOX; 1.
 CC DR PROSITE; PS00027; HOMEBOX_1; 1.
 CC DR PROSITE; PS00027; HOMEBOX_2; 1.
 CC DR Transcription regulation; DNA-binding; Homeobox; Nuclear protein;
 CC Activator; Trans-acting factor; 3D-structure.
 CC KM Transcription regulation; DNA-binding; Homeobox; Nuclear protein;
 CC Activator; Trans-acting factor; 3D-structure.
 CC FT DOMAIN 1 71 80 DIMERIZATION.
 CC FT ASP/GLU-RICH (ACIDIC) (POTENTIAL INVOLVEMENT WITH TRANSCRIPTION).
 CC FT DOMAIN 2 197 205 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DNA BIND 199 258 HOMEBOX.
 CC FT DOMAIN 238 258 21 AMINO ACID LOOP BETWEEN HELIX 2 AND 3.
 CC SQ SEQUENCE 628 AA; 67237 MW; 737920D1A36B9DD CRC64;
 CC -----
 CC Query Match 1.1%; Score 7; DB 1; Length 628;
 CC Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 344 SATSEEA 350
 CC DB 325 SATSEEA 331
 CC -----
 CC RESULT 21
 CC HNF1A RAT STANDARD; PRT; 628 AA.
 CC AC P1557; RAT
 CC DT 01-APR-1990 (Rel. 14, Created)
 CC DT 01-APR-1990 (Rel. 14, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Hepatocyte nuclear factor 1-alpha (HNF-1A) (Liver-specific
 CC transcription factor Lf-B1) (LFB1).
 CC GN TCFL1 OR HNF1A OR HNF-1A OR HNF-1.
 CC OS Rattus norvegicus (Rat).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC OX NCBI_TaxId=10116;
 CC RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=90003224; PubMed=2571419;
 RA Fraim M., Swart G., Monaci P., Nicosia A., Staemfli S., Frank R.,
 RA Cortese R.;
 RT "The liver-specific transcription factor Lf-B1 contains a highly
 RT diverged homeobox DNA binding domain.";
 RL Cell 59:145-157(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91016926; PubMed=2216777;
 RA Chouard T., Blumenfeld M., Bach I., Vandekerckhove J., Cereghini S.,
 RA Yaniv M.;
 RT "A distal dimerization domain is essential for DNA-binding by the
 RT atypical HNF1 homeodomain.";
 RL Nucleic Acids Res. 18:5853-5863(1990).
 RN [3]
 RP SEQUENCE OF 166-628 FROM N.A.
 RX MEDLINE=90249741; PubMed=1970973;
 RA Baumhueter S., Mendel D.B., Conley P.B., Kuo C.J., Turk C.,
 RA Graves M.K., Edwards C.A., Courtois G., Crabtree G.R.;
 RT "HNF-1 shares three sequence motifs with the POU domain proteins and
 RT is identical to Lf-B1 and APF.";
 RL Genes Dev. 4:372-379(1990).
 RN [4]
 RP SEQUENCE OF 1-12 FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Tomei L., Piaggio G., Tonietti C., Lazzerio D., de Francesco R.,
 RA Pozzi L., Gerstner J., Cortese R.;
 RL Submitted (AUG-1992) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP POSITION OF HOMEBOX.
 RX MEDLINE=90106643; PubMed=1967225;
 RA Finney M.;
 RT "The homeodomain of the transcription factor Lf-B1 has a 21 amino
 RT acid loop between helix 2 and helix 3.";
 RL Cell 60:5-6(1990).
 RN [6]
 RP STRUCTURE BY NMR OF 1-32.
 RX MEDLINE=91105074; PubMed=1988016;
 RA Pastore A., de Francesco R., Barbato G., Castiglione Morelli M.A.,
 RA Motta A., Cortese R.;
 RT "1H resonance assignment and secondary structure determination of the
 RT dimerization domain of transcription factor LfB1.";
 RL Biochemistry 30:148-153(1991).
 RN [7]
 RP STRUCTURE BY NMR OF 195-286.
 RX MEDLINE=93259120; PubMed=8491172;
 RA Leitling B., de Francesco R., Tomei L., Cortese R., Otting G.,
 RA Wuehrlich K.;
 RT "The three-dimensional NMR-solution structure of the polypeptide
 RT fragment 195-286 of the LfB1/HNF1 transcription factor from rat liver
 RT comprises a nonclassical homeodomain.";
 RL EMBO J. 12:1797-1803(1993).
 RN [8]
 RP STRUCTURE BY NMR OF 195-286.
 RX MEDLINE=97272000; PubMed=9126845;
 RA Schott O., Billeter M., Leitling B., Wider G., Wuehrlich K.;
 RT "The NMR solution structure of the non-classical homeodomain from the
 RT rat liver LfB1/HNF1 transcription factor.";
 RL J. Mol. Biol. 267:673-683(1997).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 195-286.
 RX MEDLINE=93259121; PubMed=8491173;
 RA Ceska T.A., Lamers M., Monaci P., Nicosia A., Cortese R., Suck D.;
 RT "The X-ray structure of an atypical homeodomain present in the rat
 RT liver transcription factor LfB1/HNF1 and implications for DNA
 RT binding.";
 RL EMBO J. 12:1805-1810(1993).
 CC -1- FUNCTION: Required for the expression of several liver specific
 CC genes. Binds to the inverted palindrome 5'-GTTAATATTAC-3'.
 CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC RN [1]

CC	-I- TISSUE SPECIFICITY: Liver.
CC	-I- SIMILARITY: Belongs to the HNF1 homebox family.
CC	-I- SIMILARITY: Contains 1 homeobox domain.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
CC	at the European Bioinformatics Institute. There are no restrictions on
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; J03170; AAA4524.1; -
DR	EMBL; X54423; CAA38295.1; -
DR	EMBL; X67649; CAA47891.1; -
DR	EMBL; X53297; CAA37387.1; ALT_INIT.
DR	PIR; A33333; A33333.
DR	PIR; S25485; S25485.
DR	PDB; 1LFB; 31-OCT-93.
DR	PDB; 2LFB; 12-MAR-97.
DR	TRANSFAC; T00369; -
DR	InterPro; IPR006899; HNF-1_N.
DR	InterPro; IPR006898; HNF1A_C.
DR	InterPro; IPR006897; HNF1B_C.
DR	InterPro; IPR001356; Homeobox.
DR	Pfam; PF04814; HNF-1_N; 1.
DR	Pfam; PF04813; HNF-1A_C; 1.
DR	Pfam; PF04812; HNF-1B_C; 1.
DR	SMART; SM00389; HOX; 1.
DR	PROSITE; PS00027; HOMEBOX_1; 1.
DR	PROSITE; PS50071; HOMEBOX_2; 1.
KW	Transcription regulation; DNA-binding; Homeobox; Nuclear protein;
KW	Activator; Trans-acting factor; 3D-structure.
FT	DOMAIN 1 31
FT	DOMAIN 71 80
FT	FT
FT	DOMAIN 197 205
FT	FT
FT	DNA_BIND 199 279
FT	FT
FT	DOMAIN 238 258
FT	FT
FT	HELI_X 208 218
FT	TURN 219 220
FT	HELI_X 226 241
FT	TURN 242 245
FT	TURN 248 249
FT	TURN 251 254
FT	HELI_X 255 257
FT	HELI_X 261 273
FT	TURN 274 275
SO	SEQUENCE 628 AA; 67213 MW; 8D28099308C6A52 CRC64;
Query Match	
Best Local Similarity 1.1%; Score 7; DB 1; Length 628;	
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Oy	344 SATSEA 350
Db	325 SATSEA 331
RESULT 22	
PDAT_YEAST	
ID	PDAT_YEAST
AC	PA0345; STANDARD; PRT; 661 AA.
DT	01-FEB-1995 (Rel. 31, Created)
DT	01-FEB-1995 (Rel. 31, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Phospholipid:diacylglycerol acyltransferase (EC 2.3.1.158) (PDAT).
GN	LR01 OR YNR008W OR N2042.
OS	Saccharomyces cerevisiae (Baker's yeast).
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_XX	NCBI_TaxId=4932;
NN	[1]

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=S288c / Fy1679;
RC      MEDLINE=95208356; PubMed=7900425;
RA      Verhaesele P., Aert R., Voet M., Volckaert G.;
RT      "Twelve open reading frames revealed in the 23.6 kb segment flanking
RT      the centromere on the Saccharomyces cerevisiae chromosome XIV right
RL      arm.";
RL      Yeast 10:1355-1361 (1994).
RN      (2)
RP      CHARACTERIZATION.
RX      MEDLINE=20300927; PubMed=10829075;
RA      Dahlqvist A., Stehl U., Lenman M., Banas A., Lee M., Sandager L.,
RT      Rome H., Stymne S.;
RT      "Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes
RT      the acyl-CoA-independent formation of triacylglycerol in yeast and
RT      plants.";
RT      Proc. Natl. Acad. Sci. U.S.A. 97:6487-6492 (2000).
CC      -1- FUNCTION: Triacylglycerol formation by an acyl-CoA independent
CC      pathway. The enzyme specifically transfers acyl groups from the
CC      sn-2 position of a phospholipid to diacylglycerol, thus forming
CC      an sn-1-lysophospholipid.
CC      -1- CATALYTIC ACTIVITY: Phospholipid + 1,2-diacylglycerol =
CC      lysophospholipid + triacylglycerol.
CC      -1- SIMILARITY: SOME. TO MAMMALIAN PHOSPHATIDYLCHOLINE-STEROL O-
CC      ACYLTRANSFERASE.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; X77395; CAAS4576.1; -.
DR      EMBL; Z71623; CAA96285.1; -.
DR      PIR; S45131; S45131.
DR      Germonline; 143353; -.
DR      SGD; S0005291; LRO1.
DR      GO; GO:0046027; F:phospholipid:diacylglycerol acyltransferase. . . ; IDA.
DR      GO; GO:0019915; P:lipid storage; IDA.
DR      GO; GO:0019432; P:triacylglycerol biosynthesis; IDA.
DR      InterPro; IPR003386; LACT.
DR      Pfam; PF02450; LACT; 1.
DR      Trnstransferase; Acyltransferase; Transmembrane.
DR      KMW TRANSMEM 81 101 POTENTIAL.
SQ      SEQUENCE 661 AA; 75393 MW; 01C043319A036F44 CRC64;
-----
Query Match 1.1%; Score 7; DB 1; Length 661;
Best Local Similarity 100.0%; Ptd. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      421 SSSSDAL 427
      |||||
      431 SSSSDAL 437

RESULT 23
V244_FOWPV
ID      V244_FOWPV STANDARD; PRT; 668 AA.
AC      O93426;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Putative ankryrin-repeat protein FPV244.
GN      FPV244.
OS      Fowlpox virus (FPV).
OC      Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC      Avipoxvirus.
OX      NCBI_TaxID=10261;
RX      MEDLINE=20193820; PubMed=10729156;

```

RA Alfonso C.L., Tulman E.R., Lu Z., Zaak L., Kutish G.F., Rock D.L.;
 RT "The genome of fowlpox virus";
 CC J. Virol. 74:3815-3831(2000).
 CC -1- SIMILARITY: Contains 12 ANK repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 DR HSMB; AP198100; AAF4568.1; -
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; ank; 10.
 DR SMART; SM00248; ANK; 11.
 DR PROSITE; PSS0088; ANK_REPEAT; 7.
 DR PROSITE; PSS0297; ANK_REPEAT_REGION; 1.
 KW Hypothetical protein; Repeat; ANK repeat.
 FT REPEAT
 FT 40 69 ANK 1.
 FT REPEAT 144 173 ANK 2.
 FT REPEAT 177 206 ANK 3.
 FT REPEAT 210 239 ANK 4.
 FT REPEAT 272 302 ANK 5.
 FT REPEAT 306 336 ANK 6.
 FT REPEAT 340 370 ANK 7.
 FT REPEAT 374 403 ANK 8.
 FT REPEAT 407 437 ANK 9.
 FT REPEAT 441 471 ANK 10.
 FT REPEAT 473 502 ANK 11.
 FT REPEAT 571 602 ANK 12.
 SQ SEQUENCE 668 AA; 76200 MW; F3C5BA076023791B CRC64;
 QY 177 KELLION 183
 Db 58 KELLION 64
 Query Match 1.1%; Score 7; DB 1; Length 668;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 24
 MYB_HUMAN STANDARD; PRT; 700 AA.
 ID MYB_HUMAN
 AC P10244;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Myb-related protein B (B-Myb).
 GN MYBL2 OR MYB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90083548; PubMed=3060855;
 RA Nomura N., Takahashi M., Matsui M., Ishii S., Date T., Sasamoto S.,
 RA Ishizaki R.;
 RT "Isolation of human cDNA clones of myb-related genes, A-myb and
 RT B-myb";
 RL Nucleic Acids Res. 16:11075-11089(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Cowell G.J., Deadman R., Dhand P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Hammond D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hargrave S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leveson-Gilho M.H., Leverhna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McInlay K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Piliavin B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showlgreen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Walli M., Wallis J.W.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, and Eye;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.E.,
 RA Brownstein M.J., Ueda T.B., Tomshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughlano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Blake J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Feley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP PHOSPHORYLATION OF THR-444; THR-487; THR-494 AND SER-577.
 RX MEDLINE=99054667; PubMed=9840932;
 RA Saville M.K., Watson R.U.;
 RT "The cell-cycle regulated transcription factor B-Myb is phosphorylated
 RT by cyclin A/Cdk2 at sites that enhance its transactivation
 RT properties";
 RL Oncogene 17:2679-2689(1998).
 RN [5]
 RP PHOSPHORYLATION OF THR-440; THR-444; THR-494; THR-520 AND SER-577.
 RX MEDLINE=99195476; PubMed=10095772;
 RA Bartsch O., Horstmann S., Toprak K., Klempner K.H., Ferrari S.;
 RT "Identification of cyclin A/Cdk2 phosphorylation sites in B-Myb";
 RL Eur. J. Biochem. 260:384-391(1999).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PTM: PHOSPHORYLATED BY CYCLIN A/CDK2 DURING S-PHASE.
 CC PHOSPHORYLATION AT THR-520 IS PROBABLY INVOLVED IN TRANSCRIPTIONAL
 CC ACTIVITY.
 CC -1- SIMILARITY: Contains 3 Myb-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).

```

CC -----
DR EMBL: AL13293; CAA31655.1; -
DR EMBL: X121866; CAC08392.1; -
DR EMBL: BC007585; AA07558.1; -
DR EMBL: BC053555; AA03355.1; -
DR PIR: S01991; S01991.
DR HSSP: Q03237; 1A5J.
DR TRANSFAC: T00065; -
DR Genew: HGNC:7548; MYBL2.
DR GK: P10244; -
DR MIM: 601415; -
DR GO: GO:0003700; F:transcription factor activity; TAS.
DR InterPro: IPR001005; Myb DNA binding.
DR Pfam: PF00249; myb DNA-binding; 3.
DR SMART: SM00717; SANT; 3.
DR PROSITE: PS00034; MYB_1; 3.
DR PROSITE: PS00334; MYB_2; 3.
DR PROSITE: PS50090; MYB_3; 3.
DR Transcription regulation; Nuclear protein; DNA-binding; Repeat;
Phosphorylation.
KW DNA_BIND 26 77 MYB 1.
FT DNA_BIND 78 129 MYB 2.
FT MOD_RES 130 180 MYB 3.
FT MOD_RES 440 440 PHOSPHORYLATION (BY CDK2).
FT MOD_RES 444 444 PHOSPHORYLATION (BY CDK2).
FT MOD_RES 487 487 PHOSPHORYLATION (BY CDK2).
FT MOD_RES 494 494 PHOSPHORYLATION (BY CDK2).
FT MOD_RES 520 520 PHOSPHORYLATION (BY CDK2).
FT MOD_RES 577 577 PHOSPHORYLATION (BY CDK2).
SQ SEQUENCE 700 AA; 78764 MW; D91B28B3DAB94061 CRC64;

```

```

Query Match 1.1%; Score 7; DB 1; Length 700;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 81 CSQKVVV 87
DB 479 CSQKVVV 485

```

```

RESULT 25
MYB_MOUSE
ID MYB_MOUSE STANDARD; PRT; 704 AA.
AC P48972;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myb-related protein B (B-Myb).
GN MYBL2 OR BMYB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9236176; PubMed=1501895;
RA Lam E.W., Robinson C., Watson R.J.;
RT "Characterization and cell cycle-regulated expression of mouse
RT B-myb.";
RL Oncogene 7:1685-1690(1992).
RP SEQUENCE FROM N.A.
RC TISSUE=Limb;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lonnellano N.A., Peters G.J., Abramson R.D., Mollany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

```

RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green S.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalske U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maita M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 1-6 FROM N.A.
RX MEDLINE=93327760; PubMed=8334989;
RA Lam E.W., Watson R.J.;
RT "An E2F-binding site mediates cell-cycle regulated repression of
RT mouse B-myb transcription.";
RL EMBO J. 12:2705-2713(1993).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: PHOSPHORYLATED BY CYCLIN A/CDK2 DURING S-PHASE.
CC PHOSPHORYLATION AT THR-524 IS PROBABLY INVOLVED IN TRANSCRIPTIONAL
CC ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: Contains 3 Myb-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: X70472; CAA49898.1; -
DR EMBL: BC050842; AA050842.1; -
DR EMBL: X73028; AA051511.1; -
DR PIR: S33704; S33704.
DR HSSP: Q03237; 1A5J.
DR MGD: MGI:101785; Myb12.
DR InterPro: IPR001005; Myb DNA binding.
DR Pfam: PF00249; myb DNA-binding; 3.
DR SMART: SM00717; SANT; 3.
DR PROSITE: PS00037; MYB_1; 3.
DR PROSITE: PS00334; MYB_2; 3.
DR PROSITE: PS50090; MYB_3; 3.
KW Transcription regulation; Nuclear protein; DNA-binding; Repeat;
Phosphorylation.
FT DNA_BIND 26 77 MYB 1.
FT DNA_BIND 78 129 MYB 2.
FT DNA_BIND 130 180 MYB 3.
FT MOD_RES 443 443 PHOSPHORYLATION (BY CDK2) (BY
FT MOD_RES 447 447 SIMILARITY).
FT MOD_RES 490 490 PHOSPHORYLATION (BY CDK2) (BY
FT MOD_RES 490 490 SIMILARITY).
FT MOD_RES 497 497 PHOSPHORYLATION (BY CDK2) (BY
FT MOD_RES 497 497 SIMILARITY).
FT MOD_RES 524 524 PHOSPHORYLATION (BY CDK2) (BY
FT MOD_RES 524 524 SIMILARITY).
FT MOD_RES 581 581 PHOSPHORYLATION (BY CDK2) (BY
FT MOD_RES 581 581 SIMILARITY).
SQ SEQUENCE 704 AA; 79102 MW; 0EF09C1EE2184E47 CRC64;

```

```

Query Match 1.1%; Score 7; DB 1; Length 704;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 81 CSQKVVV 87
DB 482 CSQKVVV 488

```

RESULT 26

```

BGAL_MALDO          STANDARD;          PRT;          731 AA.
ID_BGAL_MALDO
AC_P48981;
DT_01-FEB-1996 (Rel. 33, Created)
DT_01-FEB-1996 (Rel. 33, Last sequence update)
DT_28-FEB-2003 (Rel. 41, Last annotation update)
DE_Beta-galactosidase precursor (EC 3.2.1.23) (Lactase) (Acid beta-
DE_galactosidase) (Exo-(1-->4)-beta-D-galactanase).
OS_Malus domestica (Apple) (Malus sylvestris)
OC_Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC_Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC_eurosid1; Rosales; Rosaceae; Maloideae; Malus.
OX_NCBI_TaxID=3750;
RN_1;
RP_SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC_STRAIN=cv. Granny Smith; TISSUE=fruit cortical tissue;
RC_MEDLINE=95083752; PubMed=791682;
RA_Ross G.S., Weigryn T., Macrae E.A., Redgwell R.J.;
RT "Apple beta-galactosidase. Activity against cell wall polysaccharides
RT and characterization of a related cDNA clone.";
RL_Plant Physiol. 106:521-528(1994).
CC_-1- FUNCTION: Involved in cell wall degradation. Degrades
CC_polysaccharides containing beta-(1-->4)-linked galactans, acting
CC_as an exo-(1-->4)-beta-D-galactanase.
CC_-1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC_galactose residues in beta-D-galactosides.
CC_-1- SIMILARITY: Belongs to family 35 of glycosyl hydrolases.
CC-----
CC_This SWISS-PROT entry is copyright. It is produced through a collaboration
CC_between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC_the European Bioinformatics Institute. There are no restrictions on its
CC_use by non-profit institutions as long as its content is in no way
CC_modified and this statement is not removed. Usage by and for commercial
CC_entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC_or send an email to license@sb-sib.ch).
CC-----
CC_EMBL; L29451; AAA62324.1; -.
DR_PIR; T17002; T17002.
DR_InterPro; IPR008979; Gal_bind_1like.
DR_InterPro; IPR001944; Glyco_hydro_35.
DR_Pfam; PF01301; Glyco_hydro_35; 1.
DR_PRINITS; PR00742; GLYHDRLABE35.
DR_PROSITE; PS01182; GLYCOSYL_HYDROL_F35; 1.
KM_Hydrolase; Glycosidase; Signal.
FT_CHAIN 1 23
FT_ACT_SITE 24 731
FT_ACT_SITE 182 182
FT_ACT_SITE 251 251
FT_SEQUENCE 731 AA; 80995 MW; FAB65D24A0D30BD4 CRC64;

Query Match          1.1%; Score 7; DB 1; Length 731;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 648 KVSQSS 654
Db 414 KVSQSS 420

```

```

RP_SEQUENCE FROM N.A.
RC_STRAIN=22491 / Serogroup A / Serotype 4A;
RC_MEDLINE=20222556; PubMed=10761919;
RX_Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA_Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA_Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holtroyd S.,
RA_Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA_Randreadson M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA_Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491.";
RL_Nature 404:502-506(2000).
CC_-1- SIMILARITY: Contains 1 SI motif domain.
CC-----
CC_This SWISS-PROT entry is copyright. It is produced through a collaboration
CC_between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC_the European Bioinformatics Institute. There are no restrictions on its
CC_use by non-profit institutions as long as its content is in no way
CC_modified and this statement is not removed. Usage by and for commercial
CC_entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC_or send an email to license@sb-sib.ch).
CC-----
CC_EMBL; AL162752; CAB83508.1; -.
DR_PIR; E82013; E82013.
DR_HSSP; P05055; ISRO.
DR_InterPro; IPR008994; Nucleic_acid_OB.
DR_InterPro; IPR003029; SI.
DR_InterPro; IPR006641; YqgFc.
DR_Pfam; PF00575; SI; 1.
DR_SWART; SM00316; SI; 1.
DR_SWART; SM00732; YqgFc; 1.
DR_PROSITE; PS50126; SI; 1.
KM_Hypothetical protein; RNA-binding; Complete proteome.
FT_DOMAIN 640 709
FT_DOMAIN 265 266
FT_SEQUENCE 757 AA; 83161 MW; C7E9731BD4BC0F3 CRC64;

Query Match          1.1%; Score 7; DB 1; Length 757;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 638 SENRSRE 644
Db 723 SENRSRE 729

```

```

RESULT 27
YHGF_NEIMA          STANDARD;          PRT;          757 AA.
ID_YHGF_NEIMA
AC_P57072;
DT_16-OCT-2001 (Rel. 40, Created)
DT_16-OCT-2001 (Rel. 40, Last sequence update)
DT_16-OCT-2001 (Rel. 40, Last annotation update)
DE_Hypothetical protein NMA0194.
GN_NMA0194.
OS_Neisseria meningitidis (serogroup A).
OC_Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC_Neisseriaceae; Neisseria.
OX_NCBI_TaxID=56599;
RN_1;

```

```

RESULT 28
YHGF_NEIMB          STANDARD;          PRT;          757 AA.
ID_YHGF_NEIMB
AC_O51152;
DT_01-NOV-1997 (Rel. 35, Created)
DT_16-OCT-2001 (Rel. 40, Last sequence update)
DT_16-OCT-2001 (Rel. 40, Last annotation update)
DE_Hypothetical protein NMB0075.
GN_NMB0075.
OS_Neisseria meningitidis (serogroup B).
OC_Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC_Neisseriaceae; Neisseria.
OX_NCBI_TaxID=491;
RN_1;
RP_SEQUENCE FROM N.A.
RC_STRAIN=B1940 / Serogroup B;
RC_MEDLINE=96236055; PubMed=8655518;
RX_Petering H., Hammererschmidt S., Froesch M., van Putten J.P.M.,
RA_Ison C.A., Robertson B.D.;
RT "Genes associated with meningococcal capsule complex are also found
RT in Neisseria gonorrhoeae.";
RL_Bacteriol. 178:3342-3345(1996).
RN_1;
RN_12;
RP_SEQUENCE FROM N.A.
RC_STRAIN=MC58 / Serogroup B;
RC_MEDLINE=20175755; PubMed=10710307;
RX_Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,

```

```

RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwin M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Macon T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Yamahavan J.,
RA Gill J., Scariato V., Massimani V., Pizsa M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58";
RL Science 287:1809-1815(2000).
CC -1- SIMILARITY: Contains 1 S1 motif domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L09189; AAC37046.1; -.
DR EMBL; AE002366; AAF40542.1; -.
DR PIR; A81242; A81242.
DR HSP; P05055; JSRO.
DR TIGR; NMB0075; -.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR003029; S1.
DR Pfam; PF00575; S1; 1.
DR SMART; SM00316; S1; 1.
DR SMART; SM00732; YqgFc; 1.
DR PROSITE; PS50126; S1; 1.
DR Hypothetical protein; RNA-binding; Complete proteome.
KM DOMAIN 640 709 S1 MOTIF
FT CONFLICT 265 266 WL -> CV (IN REF. 1).
SQ SEQUENCE 757 AA; 83161 MW; 7B6851EB5ED6AE CRC64;

Query Match 1.1%; Score 7; DB 1; Length 757;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 638 SENSRE 644
DB 723 SENSRE 729

RESULT 29
DEGY CAEEL STANDARD; PRT; 795 AA.
AC 001635;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Degenerin-like protein ZK770.1 in chromosome I.
GN ZK770.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chordata; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Brictol N2.
RA Wagerl L., Gattung S., Bartko L.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the amiloride-sensitive sodium channel
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U97404; AAB33309.1; -.
DR PIR; T34468; T34468.
DR WormRep; ZK770.1; CB15411.
DR InterPro; IPR004726; Deg-1.
DR InterPro; IPR001873; Na+channel_ASC.
DR Pfam; PF00858; ASC; 1.
DR PRINTS; PR01078; AMINACHANNEL.
DR PROSITE; PS01206; ASC; 1.
DR Hypothetical protein; Ionic channel; Transmembrane; Ion transport;
KM Glycoprotein; Sodium channel.
FT DOMAIN 1 58 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 59 79 POTENTIAL.
FT DOMAIN 80 739 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 740 760 POTENTIAL.
FT DOMAIN 761 795 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 795 AA; 90776 MW; 2CAAC7D41770B54 CRC64;

Query Match 1.1%; Score 7; DB 1; Length 795;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 406 SVSYSAS 412
DB 665 SVSYSAS 671

RESULT 30
NAL6 MOUSE
ID NAL6 MOUSE STANDARD; PRT; 843 AA.
AC Q91W52; Q8K014;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE PYRIN-containing APAF1-like protein 5-1ike.
GN NALP6 OR PYPAF5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP CONCEPTUAL TRANSLATION OF 1-287.
RA Hinz U.;
RL Unpublished observations (FEB-2003).
RN [2]
RP SEQUENCE OF 288-843 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins B.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stadler M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toehiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huiy S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,

```

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
 RN [3]
 RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
 RX MEDLINE=22520874; PubMed=12633874;
 RA Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
 RT "Identification of mammalian orthologs associates PYPAF5 with distinct
 functional roles";
 RL FEBS Lett. 538:173-177(2003).
 CC -1- FUNCTION: May mediate activation of CASP1 via ASC and promote
 activation of NF-kappa-B (By similarity).
 CC -1- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Contains 1 DAPIN domain.
 CC -1- SIMILARITY: Contains 1 NACHT domain.
 CC -1- SIMILARITY: Contains 1 leucine-rich (LRR) repeats.
 CC -1- CAUTION: The N-terminus was extended using ESTs and genomic
 sequences, in analogy to ortholog sequences.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; BC013519; AAH13519.1; -.
 DR EMBL; BC011139; AAH11139.1; ALT_INIT.
 DR MGI; MGI:2141990; Pypaf5.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 1.
 DR Pfam; PF05729; NACHT; 1.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 DR PROSITE; PS50837; NACHT; 1.
 DR ATP-binding; Leucine-rich repeat; Repeat.
 DR DOMAIN; 11 102 DAPIN.
 FT DOMAIN 168 485 NACHT.
 FT DOMAIN 569 585 ASP/GU-RICH.
 FT DOMAIN 654 661 POLY-LYS.
 FT REPEAT 434 459 LRR 1.
 FT REPEAT 609 632 LRR 2.
 FT REPEAT 811 834 LRR 3.
 FT NP BIND 174 181 ATP (POTENTIAL).
 SO SEQUENCE 843 AA; 94592 MW; 35FB7A766A47DB51 CRC64;

Query Match 1.1%; Score 7; DB 1; Length 843;
 Best Local Similarity 100.0%; Pred. 1.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 AYRFVKE 276
 DB 343 AYRFVKE 349

RESULT 31
 ID NAL6_RAT STANDARD; PRT; 854 AA.
 AC Q63035;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE PRIN-containing ARAFI-like protein 5-like (Angiotensin II/vasopressin
 receptor).
 GN NALP6 OR PYPAF5 OR AVR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RP CONCEPTUAL TRANSLATION OF 1-343.
 RA Hinz U.;
 RL Unpublished observations (FEB-2003).
 RN [2]
 RP SEQUENCE OF 344-854 FROM N.A., MUTAGENESIS OF GLU-393, AND FUNCTION.
 RC TISSUE=Kidney;
 RX MEDLINE=96071640; PubMed=7489366;
 RA Ruiz-Opazo N., Akimoto K., Herrera V.L.M.;
 RT "Identification of a novel dual angiotensin II/vasopressin receptor on
 the basis of molecular recognition theory";
 RL Nat. Med. 1:1074-1081(1995).
 RN [3]
 RP VARIANTS SALT-SENSITIVE HYPERTENSION SER-492 AND ARG-536.
 RX MEDLINE=21980185; PubMed=11984003;
 RA Ruiz-Opazo N., Lopez L.V., Herrera V.L.M.;
 RT "The dual AngII/AVP receptor gene N19S/C163R variant exhibits
 sodium-induced dysfunction and cosegregates with salt-sensitive
 hypertension in the Dahl salt-sensitive hypertensive rat model";
 RL Mol. Med. 8:24-32(2002).
 RN [4]
 RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
 RX MEDLINE=22520874; PubMed=12633874;
 RA Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
 RT "Identification of mammalian orthologs associates PYPAF5 with distinct
 functional roles";
 RL FEBS Lett. 538:173-177(2003).
 CC -1- FUNCTION: May mediate activation of CASP1 via ASC and promote
 activation of NF-kappa-B (By similarity). Angiotensin II and
 vasopressin binding protein. May stimulate cAMP accumulation.
 CC -1- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- TISSUE SPECIFICITY: Highly expressed in kidney. Detected at low
 levels in all tissues tested.
 CC -1- DISEASE: Defects in NALP6 may be a cause of salt-sensitive
 hypertension.
 CC -1- SIMILARITY: Contains 1 DAPIN domain.
 CC -1- SIMILARITY: Contains 1 NACHT domain.
 CC -1- SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
 CC -1- CAUTION: The N-terminus was extended using ESTs and genomic
 sequences, in analogy to ortholog sequences.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M85183; AAA03623.1; ALT_INIT.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF05729; NACHT; 1.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 DR PROSITE; PS50837; NACHT; 1.
 DR ATP-binding; Leucine-rich repeat; Repeat.
 DR DOMAIN; 1 103 DAPIN.
 FT DOMAIN 168 484 NACHT.
 FT DOMAIN 564 604 ASP/GU-RICH.
 FT DOMAIN 655 662 POLY-LYS.
 FT REPEAT 433 458 LRR 1.
 FT REPEAT 610 633 LRR 2.
 FT REPEAT 722 745 LRR 3.
 FT REPEAT 812 836 LRR 4.
 FT NP BIND 174 181 ATP (POTENTIAL).
 FT VARIANT 492 492 N -> S (IN SALT-SENSITIVE HYPERTENSION).

```

FT VARIANT 536 536 C -> R (IN SALT-SENSITIVE HYPERTENSION).
FT MUTAGEN 766 766 E->K: ABOLISHES ANGIOTENSIN II BINDING.
SQ SEQUENCE 854 AA; 95292 MW; D7BBD922D77B3734 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 854;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 270 AYPFKE 276
Db 343 AYPFKE 349

RESULT 32
NAL6_HUMAN
ID NAL6_HUMAN STANDARD; PRT; 892 AA.
AC P59044;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE NACHT, LRR- and PYD-containing protein 6 (PYRIN-containing APAF1-like
protein 5)
GN NALP6 OR PYPAF5.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22162427; PubMed=12019269;
RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
Lora J.M., Geddes B.J., Briskin M., Distefano P.S., Bertin J.;
"PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
activation of NF-kappa B and caspase-1-dependent cytokine
processing.";
RT J. Biol. Chem. 277:29674-29680(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22451042; PubMed=12563287;
RA Tschopp J., Martinon F., Burns K.;
"NALP6: a novel protein family involved in inflammation.";
RT Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
RN [3]
RP FUNCTION.
RX MEDLINE=22275822; PubMed=12387869;
RA Grenier J.M., Wang L., Manji G.A., Huang W.-J., Al-Garawi A.,
Kelly R., Carlson A., Merriam S., Lora J.M., Briskin M.,
Distefano P.S., Bertin J.;
"Functional screening of five PYPAF family members identifies PYPAFs
as a novel regulator of NF-kappa B and caspase-1.";
RT FEBS Lett. 530:73-78(2002).
CC -1- FUNCTION: May mediate activation of CASP1 via ASC and promote
activation of NF-kappa-B.
CC -1- SUBUNIT: Binds to ASC with its DAPIN domain.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: Highly expressed in granulocytes. Detected at
much lower levels in T-cells.
CC -1- SIMILARITY: Contains 1 DAPIN domain.
CC -1- SIMILARITY: Contains 1 NACHT domain.
CC -1- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL, AF479748; AAL87105.1; -.
CC EMBL, AY154461; AAO18157.1; -.
CC Genew; HGNC:22944; NALP6.
CC InterPro; IPR000767; Disease_resist.

```

```

DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF05729; NACHT; 1.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PRINTS; PR00364; DISFASERISIT.
DR SMART; SM00368; LRR_R1; 3.
DR PROSITE; PS0824; DAPIN; 1.
DR PROSITE; PS0837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 1 103 DAPIN.
FT DOMAIN 196 513 NACHT.
FT DOMAIN 604 614 POLY-GLU.
FT REPEAT 462 487 LRR 1.
FT REPEAT 727 747 LRR 2.
FT REPEAT 755 778 LRR 3.
FT REPEAT 811 834 LRR 4.
FT REPEAT 845 868 LRR 5.
FT NP_BIND 202 209 ATP (POTENTIAL).
SQ SEQUENCE 892 AA; 98733 MW; 4AABDFC76DDE9D CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 892;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 270 AYPFKE 276
Db 371 AYPFKE 377

RESULT 33
TRG1_ECOLI
ID TRG1_ECOLI STANDARD; PRT; 938 AA.
AC P33790;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trg protein.
GN TRG.
OS Escherichia coli.
OG Plasmid F.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RC MEDLINE=92204127; PubMed=1348105;
RA Firth N., Skurray R.A.;
"Characterization of the F plasmid bifunctional conjugation gene,
trg.";
RT Mol. Gen. Genet. 232:145-153(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94359430; PubMed=7915817;
RA Frost L.S., Ippen-Ihler K., Skurray R.A.;
"Analysis of the sequence and gene products of the transfer region of
the F sex factor.";
RT Microbiol. Rev. 58:162-210(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / CRE3;
RA Shimizu H., Satoh Y., Suda Y., Uehara K., Sampa G., Mizobuchi K.;
"Complete nucleotide sequence of the F plasmid: its implications for
organization and diversification of plasmid genomes.";
RT Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLAYS A CRUCIAL ROLE IN DONOR-RECIPIENT CELL-
INTERACTIONS. REQUIRED FOR TWO STAGES OF THE CONJUGATION PROCESS:
PLUS BIOSYNTHESIS AND MATING AGGREGATE STABILIZATION. MAY
INTERACT WITH TRAN.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- PTM: TRAG* PROBABLY ARISES FROM THE POST-TRANSLATIONAL CLEAVAGE OF

```



```

CC TRAG.
CC -1- CAUTION: TRAG IS NOT RESPONSIBLE FOR THE N-TERMINAL ACETYLATION
CC OF P PIIIN AS STATED BY SOME AUTHORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
CC EMBL: MS9763; AAA98081.1; -.
CC EMBL: U01159; AAC44184.1; -.
CC EMBL: AP001918; BAA97969.1; -.
CC PIR: S20480; S20480.
CC DR Ecogen; EG40114; trag.
CC KM Paenid; Conjugation; Transmembrane; Inner membrane;
CC Complete proteome.
CC FT CHAIN 1 938 TRAG PROTEIN.
CC FT DOMAIN ? 938 TRAG* PROTEIN.
CC FT TRANSSEM 1 53 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 54 73 POTENTIAL.
CC FT TRANSSEM 74 329 PERIPLASMIC (POTENTIAL).
CC FT DOMAIN 330 350 POTENTIAL.
CC FT TRANSSEM 351 412 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 413 433 POTENTIAL.
CC FT TRANSSEM 434 938 PERIPLASMIC (POTENTIAL).
CC SQ SEQUENCE 938 AA; 102471 MW; 9E09C8402ACDFFDF CRC64;

Query Match 1.1%; Score 7; DB 1; Length 938;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 417 LHGFSSS 423
| | | | |
Db 544 LHGFSSS 550

RESULT 34
AMPE RAT STANDARD; PRT; 945 AA.
AC P50123; 064200; 09JUL07; 09JUL09;
DT 01-OCT-1996 (Rel. 34, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glutamy1 aminopeptidase (EC 3.4.11.7) (EAP) (Aminopeptidase A) (APA).
GN ENPEP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND TISSUE SPECIFICITY.
RC STRAIN=Sprague-Dawley; TISSUE=Hippocampus;
RA MEDLINE=20455312; PubMed=10978538;
RX Lee H.-U., Tomioka M., Takaki Y., Masumoto H., Saito T.C.;
RT "Molecular cloning and expression of aminopeptidase A isoforms from
RT rat hippocampus."
RL Biochim. Biophys. Acta 1493:273-278(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RC TISSUE=Kidney;
RX MEDLINE=20131186; PubMed=10666143;
RA Troyanovskaya M., Jayaraman G., Song L., Healy D.P.;
RT "Aminopeptidase-A. I. cDNA cloning and expression and localization in
RT rat tissues."
RL Am. J. Physiol. 278:R413-R424(2000).
RN [3]
RP SEQUENCE OF 265-397 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95029920; PubMed=7943354;
RA Song L., Ye M., Troyanovskaya M., Walk E., Walk S., Healy D.P.;

```

```

RT "Rat kidney glutamy1 aminopeptidase (aminopeptidase A): molecular
RT identity and cellular localization."
RT Am. J. Physiol. 267:F546-F557(1994).
RL [4]
RN (4)
RP SEQUENCE OF 482-606 FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RC TISSUE=Kidney;
RX MEDLINE=96188975; PubMed=8613196;
RA Troyanovskaya M., Song L., Jayaraman G., Healy D.P.;
RT "Expression of aminopeptidase A, an angiotensinase, in glomerular
RT mesangial cells."
RL Hypertension 27:518-522(1996).
CC -1- FUNCTION: Appears to have a role in the catabolic pathway of the
CC renin-angiotensin system. Isoform 1 has aminopeptidase activity
CC while isoform 2 does not.
CC -1- CATALYTIC ACTIVITY: Release of a N-terminal glutamate (and to a
CC lesser extent aspartate) from a peptide.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=APAL;
CC IsoId=P50123-1; Sequence=displayed;
CC Name=2; Synonyms=APAS;
CC IsoId=P50123-2; Sequence=VSP_007844, VSP_007845;
CC -1- TISSUE SPECIFICITY: Highest expression in kidney proximal tubules
CC and ileum enterocytes. High expression also detected in liver and
CC pituitary. Lower levels in heart, adrenal gland and brain. Not
CC detected in aorta, lung or spleen. In heart, higher levels in
CC ventricle than in atrium. Also expressed in glomerular mesangial
CC cells.
CC -1- SIMILARITY: Belongs to peptidase family M1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
CC EMBL: AF146044; AAF66704.1; -.
CC EMBL: AF146518; AAF66710.1; -.
CC EMBL: AF214568; AAF37622.1; -.
CC MEROPS: M01.003; -.
CC DR InterPro: IPR006025; Pept_M_Zn_BS.
CC DR InterPro: IPR001930; Peptidase_M1.
CC DR Pfam: PF01433; Peptidase_M1; 1.
CC DR PRINTS: PR00756; ALADIPRTSE.
CC DR PROSITE: PS00142; ZINC_PROTEASE; 1.
CC KW Hydroxylase; Metalloprotease; Aminopeptidase; Zinc; Calcium;
CC Glycoprotein; Transmembrane; Phosphorylation; Signal-anchor;
CC Alternative splicing.
CC FT DOMAIN 1 17 CYTOPLASMIC (POTENTIAL).
CC FT TRANSSEM 18 40 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).
CC FT DOMAIN 41 945 EXTRACELLULAR (POTENTIAL).
CC FT METAL 385 385 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT ACT SITE 386 386 BY SIMILARITY.
CC FT METAL 389 389 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT METAL 408 408 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT ACT_SITE 471 471 PROTON DONOR (POTENTIAL).
CC FT MOD_RES 112 12 PHOSPHORYLATION (BY SIMILARITY).
CC FT CARBOHYD 116 116 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 189 189 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 236 236 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 316 316 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 346 346 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 584 584 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 601 601 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 640 640 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 669 669 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 754 754 N-LINKED (GLCNAC...) (POTENTIAL).

```

```

FT CARBOHYD 766 766 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 792 792 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 569 573 YTMNI -> NHREC (in isoform 2).
FT VARSPLIC 574 945 /FTId=VSP_007844.
FT VARSPLIC 574 945 Missing (in isoform 2).
FT CONFLICT 256 256 K -> E (IN REF. 1).
FT CONFLICT 306 306 N -> K (IN REF. 1).
FT CONFLICT 748 748 A -> D (IN REF. 1).
SQ SEQUENCE 945 AA; 107994 MW; 5A74F1A537DC5937 CRC64;

Query Match 1.1%; Score 7; DB 1; Length 945;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 290 QLLDYER 296
Db 660 QLLDYER 666

RESULT 35
RHG4_HUMAN STANDARD; PRT; 946 AA.
AC P98171; Q14144;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Rho-GTPase-activating protein 4 (Rho-GAP hematopoietic protein C1)
DE (p115).
GN ARHGAP4 OR RHOGAP4 OR RGCT OR KIAA0131.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA TRIBIO1 C., Dreoto S., Bione S., Cesareni G., Torrist M.R.,
RA Lotti L.V., Lanfranccone L., Toniolo D., Pellicci P.;
RA "An X chromosome-linked gene encoding a protein with characteristics
RT of a rhogap predominantly expressed in hematopoietic cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:695-699(1996).
RN [2]
RP SEQUENCE OF 5-946 FROM N.A.
RX MEDLINE=96127530; PubMed=8590280;
RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
RA "Prediction of the coding sequences of unidentified human genes. IV.
RT The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RL DNA Res. 2:167-174(1995).
RN [3]
RP SEQUENCE OF 1-103 FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=95072568; PubMed=7981673;
RA Tribio1 C., Mancini M., Plassart E., Bione S., Rivella S.,
RA Sala C., Torri G., Toniolo D.;
RA "Isolation of new genes in distal Xq28: transcriptional map and
RT identification of a human homologue of the ARD1 N-acetyl transferase
RL of Saccharomyces cerevisiae.";
Hum Mol. Genet. 3:1061-1068(1994).
CC -1- FUNCTION: Inhibitory effect on stress fiber organization. May
CC down-regulate Rho-like GTPase in hematopoietic cells.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, JUST BELOW THE PLASMA MEMBRANE.
CC -1- TISSUE SPECIFICITY: Predominantly in hematopoietic cells (spleen,
CC thymus and leukocytes); low levels in placenta, lung and various
CC fetal tissues.
CC -1- SIMILARITY: Contains 1 FCH domain.
CC -1- SIMILARITY: Contains 1 Rho-GAP domain.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X78817; CAA55394.1; -
CC EMBL; D50921; BAA09480.1; -
CC PIR; I38100; I38100.
CC HSSP; Q60631; IGBR.
CC Genew; HGNC:674; ARHGAP4.
CC MIM; 300023; -
CC GO; GO:0005737; C:cytoplasm; TAS.
CC GO; GO:0005100; F:Rho GTPase activator activity; TAS.
CC GO; GO:0005070; F:SH/SH2 adaptor protein activity; TAS.
CC GO; GO:0007011; F:regulation of cytoskeleton; TAS.
CC GO; GO:0007266; P:Rho protein signal transduction; TAS.
CC InterPro; IPR001060; Cdc15_Fes_CIP4.
CC InterPro; IPR008936; Rho GAP.
CC InterPro; IPR00198; Rho GAP.
CC InterPro; IPR01452; SH3.
CC Pfam; PF00611; FCH; 1.
CC Pfam; PF00620; RHOGAP; 1.
CC Pfam; PF00018; SH3; 1.
CC ProDom; PD000066; SH3; 1.
CC SMART; SM00055; FCH; 1.
CC SMART; SM00324; Rhogap; 1.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS50133; FCH; 1.
CC PROSITE; PS50238; RHOGAP; 1.
CC PROSITE; PS50002; SH3; 1.
CC GTPase activation; SH3 domain; Coiled coil.
FT DOMAIN 15 88 FCH
FT DOMAIN 128 195 COILED COIL (POTENTIAL).
FT DOMAIN 507 695 RHO-GAP.
FT DOMAIN 746 805 SH3.
FT CONFLICT 609 609 D -> A (IN REF. 2).
FT CONFLICT 731 731 D -> E (IN REF. 2).
SQ SEQUENCE 946 AA; 105055 MW; 8BB291410BCA7DCA CRC64;

Query Match 1.1%; Score 7; DB 1; Length 946;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 498 SLSPLH 504
Db 95 SLSPLH 101

RESULT 36
MLB_MYCTU STANDARD; PRT; 966 AA.
AC OS3653;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative membrane protein mmp11.
GN Mmp11 OR RV0202C OR Wt0212 OR Mtv033.10C OR MB0208C.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CX NCBI_TaxId=1773; 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jags K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

```

RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RN Nature 393:537-544(1998).
 [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwin M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Mayhew L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL J. Bacteriol. 184:5479-5490(2002).
 [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.bovis; STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972;
 RA Garner T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsenpe C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the mmpL family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; AL021928; CA117324.1; -;
 DR EMBL; AB006930; AAK4433.1; -;
 DR EMBL; BX248334; CAD93072.1; -;
 DR PIR; G70838; G70838.
 DR TIGR; MT0212; -;
 DR Truncation: Rv0202c; -;
 DR InterPro: IPR000731; SSD_5TM.
 DR PROSITE; PSS0156; SSD_1.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 188 208 POTENTIAL.
 FT TRANSMEM 214 234 POTENTIAL.
 FT TRANSMEM 235 255 POTENTIAL.
 FT TRANSMEM 279 299 POTENTIAL.
 FT TRANSMEM 311 331 POTENTIAL.
 FT TRANSMEM 373 393 POTENTIAL.
 FT TRANSMEM 409 429 POTENTIAL.
 FT TRANSMEM 527 547 POTENTIAL.
 FT TRANSMEM 557 577 POTENTIAL.
 FT TRANSMEM 595 615 POTENTIAL.
 FT TRANSMEM 646 666 POTENTIAL.
 FT TRANSMEM 668 688 POTENTIAL.
 SQ SEQUENCE 966 AA; 103502 MW; 3E8BF0327CBEA2DA CRC64;

Query Match 1.1%; Score 7; DB 1; Length 966;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 351 GORPVP 357
 Db 798 GORPVP 804

RESULT 37
 KPI1_MOUSE

ID KPI1_MOUSE STANDARD; PRT; 1038 AA.
 AC 099PFB;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Kinesin-like protein Kif17 (MmKif17).
 GN KIF17.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20307907; PubMed=10846156;
 RA Setou M., Nakagawa T., Seog D.-H., Hirokawa N.;
 RT "Kinesin superfamily motor protein Kif17 and mLin-10 in NMDA
 RT receptor-containing vesicle transport.";
 RL Science 288:1796-1802(2000).
 CC -1- FUNCTION: Transports vesicles containing N-methyl-D-aspartate
 CC (NMDA) receptor 2B along microtubules.
 CC -1- SUBUNIT: Interacts with LIN-10 PDZ domain.
 CC -1- TISSUE SPECIFICITY: Neuronal-specific.
 CC -1- SIMILARITY: Belongs to the kinesin-like protein family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; AB008667; BAB21099.1; -;
 DR HSSP; P17119; 3KAR.
 DR MGD; MGI:1098229; Kif17.
 DR GO; GO:0005871; C:kinesin complex; IDA.
 DR GO; GO:0003777; F:microtubule motor activity; IDA.
 DR GO; GO:0005515; F:protein binding; IPT.
 DR GO; GO:0007017; P:mitochondrion-based process; IDA.
 DR GO; GO:0016192; P:vesicle-mediated transport; IDA.
 DR InterPro: IPR001752; kinesin_motor.
 DR Pfam; PF00225; kinesin_1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE; PS00677; KINESIN_MOTOR_DOMAIN2; 1.
 KW Motor protein; Microtubule; ATP-binding; Coiled coil; Transport;
 KW Protein transport.
 FT DOMAIN 1 265 KINESIN-MOTOR.
 FT DOMAIN 346 470 COILED COIL (POTENTIAL).
 FT DOMAIN 748 855 COILED COIL (POTENTIAL).
 FT NP BIND 91 98 ATP (POTENTIAL).
 SQ SEQUENCE 1038 AA; 116372 MW; 2BED852A3AFBD46 CRC64;

Query Match 1.1%; Score 7; DB 1; Length 1038;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 343 DSATSEA 349
 Db 547 DSATSEA 553

RESULT 38
 UBP7_YEAST STANDARD; PRT; 1071 AA.
 AC P404E3;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ubiquitin carboxyl-terminal hydrolase 7 (EC 3.1.2.15) (Ubiquitin
 DE thioesterase 7) (Ubiquitin-specific processing protease 7)

```

DE (Deubiquitinating enzyme 7).
GN UBP7 OR YII156W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313266; PubMed=9169870;
RA Churcher C.M., Bowman S., Baddock K., Bankier A., Brown D.,
RA Hillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,
RA Moulé S., Odeh C., Pearson D., Rajadream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.,
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX."
RL Nature 337:84-87(1997).
CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC ubiquitin + a thiol.
CC -1- SIMILARITY: Belongs to peptidase family C19.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, Z38059; CAAB6122.1; -.
DR PIR, S48378; S48378.
DR GeneOnline, 139691; -.
DR MEROPS, C19, UPW; -.
DR SGD, S0001418; UBP7.
DR GO: GO:0005737; C:cytoplasm; IC.
DR InterPro, IPR001394; Peptidase_C19.
DR InterPro, IPR001763; Rhodanese-like.
DR Pfam, PF00581; Rhodanese; 1.
DR SMART, SM00450; RHOD; 1.
DR PROSITE, PS00972; UCH 2_1; 1.
DR PROSITE, PS00973; UCH 2_2; 1.
DR PROSITE, PS00973; UCH 2_3; 1.
DR PROSITE, PS0235; UCH 2_3; 1.
KW Ubl conjugation pathway; Hydrolase; thiol protease; Multigene family.
FT ACT SITE 618 BY SIMILARITY.
FT ACT SITE 1006 BY SIMILARITY.
FT ACT SITE 1014 BY SIMILARITY.
FT ACT SITE 1014 BY SIMILARITY.
SQ SEQUENCE 1071 AA; 123133 MW; 82683A01063CEC74 CRC64;

Query Match 1.1%; Score 7; DB 1; Length 1071;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 EXVLLD 30
Db 330 EXVLLD 336

RESULT 39
CYA7_MOUSE STANDARD; PRT; 1099 AA.
ID PS1829;
AC PS1829;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Adenylate cyclase, type VII (EC 4.6.1.1) (ATP pyrophosphate-lyase)
DE (Adenylate cyclase).
DE ADCT17.
GN ADCT17.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoma;
RX MEDLINE=95050703; PubMed=7961850;
RA Watson P.A., Krupinski J., Kempinski A.M., Frankenfeld C.D.;
RT "Molecular cloning and characterization of the type VII isoform of
RT mammalian adenylate cyclase expressed widely in mouse tissues and in
RT S49 mouse lymphoma cells."
RL J. Biol. Chem. 269:28893-28898(1994).
CC -1- FUNCTION: This is a membrane-bound, calcium-inhibitable adenylate
CC cyclase.
CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -1- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Most abundant in heart, spleen and lung.
CC -1- DOMAIN: Composed of two homologous domains.
CC -1- SIMILARITY: Belongs to the adenylate cyclase class-4/guanylyl
CC cyclase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, U12919; AAAS7554.1; -.
DR PIR, A55405; A55405.
DR HSP, P26769; IAB8.
DR MGD, MGI:102891; Adcy7.
DR InterPro, IPR01054; G_cyclase.
DR Pfam, PF00211; guanylate_cyc; 2.
DR SMART, SM00044; CYC; 2.
DR PROSITE, PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE, PS00452; GUANYLATE_CYCLASES_2; 2.
KW Lyase; GMP biosynthesis; Transmembrane; Glycoprotein; Repeat;
KW Metal-binding; Magnesium.
FT DOMAIN 1 33 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 95 117 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT DOMAIN 199 595 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 596 616 POTENTIAL.
FT TRANSMEM 621 641 POTENTIAL.
FT TRANSMEM 670 689 POTENTIAL.
FT TRANSMEM 719 738 POTENTIAL.
FT TRANSMEM 747 766 POTENTIAL.
FT TRANSMEM 813 833 POTENTIAL.
FT DOMAIN 834 1099 POTENTIAL.
FT TRANSMEM 834 1099 POTENTIAL.
FT METAL 286 286 MAGNESIUM 1 AND 2 (BY SIMILARITY).
FT METAL 287 287 MAGNESIUM 2 (VIA CARBOXYL OXYGEN) (BY
FT METAL SIMILARITY).
FT METAL 330 330 MAGNESIUM 1 AND 2 (BY SIMILARITY).
FT CARBOHYD 702 702 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1099 AA; 122767 MW; ADF3B3559CD11B96 CRC64;

Query Match 1.1%; Score 7; DB 1; Length 1099;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 EFSRCFP 133
Db 644 EFSRCFP 650

RESULT 40
LEPR_MACMU STANDARD; PRT; 1163 AA.
ID LEPR_MACMU
AC Q9MYI0; Q9MYK9; Q9MYL1; Q9MYL2;
DT 28-FEB-2003 (Rel. 41, Created)

```

DT 28-FEB-2003 (Rel. 41, last sequence update)
 DT 15-MAR-2004 (Rel. 43, last annotation update)
 DE Leptin receptor precursor (LEP-R) (OB receptor) (OB-R).
 GN LEP-R OR OBR.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
 RC TISSUE=Adipose tissue;
 RX MEDLINE=98408931; PubMed=9738551;
 RA Hotta K., Gustafson T.A., Ortmeier H.K., Bodkin N.L., Hansen B.C.;
 RT "Monkey leptin receptor mRNA: sequence, tissue distribution, and mRNA
 expression in the adipose tissue of normal, hyperinsulinemic, and type
 2 diabetic rhesus monkeys.";
 RL Obed. Res. 6:353-360(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
 RC TISSUE=Adipose tissue;
 RA Hotta K., Gustafson T.A., Ortmeier H.K., Bodkin N.L., Hansen B.C.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Gustafson T.A., Ortmeier H.K., Bodkin N.L., Hansen B.C.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Receptor for Obesity Factor (leptin). Involved in the
 regulation of fat metabolism and in a hematopoietic pathway
 regulated for normal lymphopoiesis. May play a role in
 reproduction.
 CC -1- FUNCTION: The short form (isoform A) may act to transport leptin
 to the cerebrospinal fluid (by similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=B;
 CC IsoId=Q9MYL0-1; Sequence=Displayed;
 CC Name=A;
 CC IsoId=Q9MYL0-2; Sequence=VSP_001695, VSP_001696;
 CC -1- TISSUE SPECIFICITY: Widely expressed. High expression of isoform B
 in liver, adipose tissue, hypothalamus and choroid plexus.
 CC -1- DOMAIN: The cytoplasmic domain may be essential for intracellular
 signal transduction by activation of JAK tyrosine kinase and
 STATs.
 CC -1- PTM: Phosphorylated on two tyrosine residues. Tyr-984 may be the
 major site of phosphorylation. Phosphorylation on both sites is
 required for full activity (by similarity).
 CC -1- SIMILARITY: Belongs to the type I cytokine family of receptors.
 CC Subfamily 2.
 CC -1- SIMILARITY: Contains 3 fibronectin type III domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF225874; AAF5388.1; -;
 CC EMBL: AF225875; AAF5389.1; -;
 CC EMBL: AF225873; AAF5387.1; -;
 CC EMBL: AF222960; AAF34683.1; -;
 CC HSSP: P16471; 1BP3.
 DR InterPro: IPR002996; CRIA.
 DR InterPro: IPR008957; FN_III-like.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003529; Hemtopoptn_L_F2.
 DR InterPro: IPR003531; Hemtopoptn_S_F1.
 DR Pfam: PF00041; Fn3_2.
 DR SMART: SMO0060; FN3_4.
 DR PROSITE: PS01353; HEMATOPO_REC_L_F2; 1.

DR PROSITE: PS01353; HEMATOPO_REC_S_F1; 1.
 KW Obesity; Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
 KM Alternative splicing; Phosphorylation.
 FT SIGNAL 1 21
 FT CHAIN 22 163
 FT DOMAIN 22 837
 FT TRANSMEM 838 860
 FT DOMAIN 861 1163
 FT DOMAIN 235 318
 FT DOMAIN 535 621
 FT DOMAIN 736 821
 FT DISULFID 37 88
 FT DISULFID 87 97
 FT DISULFID 129 140
 FT DISULFID 184 194
 FT DISULFID 350 410
 FT DISULFID 411 416
 FT DISULFID 471 526
 FT DISULFID 486 496
 FT MOD_RES 984 984
 FT MOD_RES 1139 1139
 FT VARSPPLIC 890 894
 FT VARSPPLIC 895 1163
 FT FT
 FT CONFLICT 889 889
 FT FT
 SQ SEQUENCE 1163 AA; 132295 MW; 687889108F851895 CRC64;
 Query Match 1.1%; Score 7; DB 1; Length 1163;
 Best local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 371 LEDSPLY 377
 DB 167 LEDSPLY 173
 RESULT 41
 LEP-R HUMAN
 ID LEP-R HUMAN STANDARD; PRT; 1165 AA.
 AC P48357; Q13592; Q13593; Q13594; Q92919; Q92920; Q92921;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Leptin receptor precursor (LEP-R) (OB receptor) (OB-R) (HUB219).
 GN LEP-R OR OBR OR DB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS B AND E).
 RC TISSUE=Brain;
 RX MEDLINE=96128129; PubMed=8548812;
 RA Tartegeia L.A., Dembeki M., Weng X., Deng N., Chleppert J.,
 RA Devos R., Richards G.J., Campfield L.A., Clark F.T., Deeds J.,
 RA Muir C., Sanket S., Moriarty A., Moore K.J., Smutko J.S.,
 RA Mays G.G., Wolf B.A., Monroe C.A., Tepper R.I.;
 RT "Identification and expression cloning of a leptin receptor, OB-R.";
 RL Cell 83:1263-1271(1995).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS A, B AND C).
 RC TISSUE=Fetal liver;
 RX MEDLINE=96398968; PubMed=8805376;
 RA Bennett B.D., Solar G.P., Yuan J.Q., Mathias J., Thomas G.R.,
 RA Matthews W.;
 RT "A role for leptin and its cognate receptor in hematopoiesis.";
 RL Curr. Biol. 6:1170-1180(1996).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM B), AND VARIANTS ARG-109 AND ARG-223.
 RX MEDLINE=97301763; PubMed=9158141;
 RA Thompson D.B., Ravussin E., Bennett P.H., Bogardus C.;

RT "Structure and sequence variation at the human leptin receptor gene in
 RT lean and obese Pima Indians."
 RL Hum. Mol. Genet. 6:675-679 (1997).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM A).
 RX MEDLINE=97215244; PubMed=9061609;
 RA Lioch S.-M., Di Marco F., Levin N., Armanini M., Xie M.H., Nelson C.,
 RA Bennett G.L., Williams M., Spencer S.A., Gurney A., de Sauvage F.J.;
 RT "Cloning and characterization of a human leptin receptor using a
 RT biologically active leptin immunoadhesin."
 RL J. Mol. Endocrinol. 18:77-85 (1997).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS A; C AND D).
 RC TISSUE=Fetal liver;
 RX MEDLINE=96206286; PubMed=8616721;
 RA Cioffi J.A., Shafer A.W., Zupancic T.J., Smith-Gbur J., Mikhail A.,
 RA Platicka D., Snodgrass H.R.;
 RT "Novel B219/OB receptor isoforms: possible role of leptin in
 RT hematopoiesis and reproduction."
 RL Nat. Med. 2:585-589 (1996).
 RN [6]
 RP ALTERNATIVE SPLICING DUE TO AN ENDOGENOUS RETROVIRUS.
 RX MEDLINE=99128223; PubMed=9929394;
 RA Kapitonov V.V., Uirka J.;
 RT "The long terminal repeat of an endogenous retrovirus induces
 RT alternative splicing and encodes an additional carboxy-terminal
 RT sequence in the human leptin receptor."
 RL J. Mol. Evol. 48:248-251 (1999).
 RN [7]
 RP CARBOHYDRATE-LINKAGE SITES, DISULFIDE BONDS, AND PARTIAL SEQUENCE.
 RX MEDLINE=99003211; PubMed=9786864;
 RA Hanin M., Arakawa T., Bures E.J., Young Y., Hui J.O., Rohde M.F.,
 RA Weicher A.A., Horan T.;
 RT "Human leptin receptor. Determination of disulfide structure and
 RT N-glycosylation sites of the extracellular domain."
 RL J. Biol. Chem. 273:28691-28699 (1998).
 RN [8]
 RP VARIANT ARG-223.
 RX MEDLINE=96270489; PubMed=8666155;
 RA Considine R.V., Considine E.L., Williams C.J., Hyde T.M., Caro J.F.;
 RT "The hypothalamic leptin receptor in humans: identification of
 RT incidental sequence polymorphisms and absence of the db/db mouse and
 RT fa/fa rat mutations."
 RL Diabetes 45:992-994 (1996).
 RN [9]
 RP VARIANTS ARG-109; ARG-204; ARG-223 AND ASN-656.
 RX MEDLINE=97289527; PubMed=9144432;
 RA Echalwid S.M., Soerensen T.D., Soerensen T.I., Tybjæerg-Hansen A.,
 RA Andersen T., Chung W.K., Leibel R.L., Pedersen O.;
 RT "Amino acid variants in the human leptin receptor: lack of association
 RT to juvenile onset obesity."
 RL Biochem. Biophys. Res. Commun. 233:248-252 (1997).
 RN [10]
 RP VARIANTS ARG-109; ARG-223 AND ASN-656.
 RX MEDLINE=97431549; PubMed=9287054;
 RA Chung W.K., Power-Kehoe L., Chua M., Chu F., Aronne L., Huma Z.,
 RA Sothorn M., Udall J.N., Kahle B., Leibel R.L.;
 RT "Exonic and intronic sequence variation in the human leptin receptor
 RT gene (LEPR)."
 RL Diabetes 46:1509-1511 (1997).
 RN [11]
 RP VARIANTS ARG-109; ARG-223; ASN-656 AND THR-675.
 RX MEDLINE=99075638; PubMed=9860295;
 RA Roth H., Korn T., Rosenkranz K., Hinney A., Ziegler A., Kunz J.,
 RA Siegfried W., Mayer H., Hebebrand J., Grzeschik K.-H.;
 RT "Transmission disequilibrium and sequence variants at the leptin
 RT receptor gene in extremely obese German children and adolescents."
 RL Hum. Genet. 103:540-546 (1998).
 RN [12]
 RP VARIANTS ARG-109; ARG-223 AND ASN-656.
 RX MEDLINE=97318795; PubMed=9175732;
 RA Gotoda T., Manning B.S., Goldstone A.P., Imrie H., Evans A.L.,
 RA Strosberg A.D., McKeigue P.M., Scott J., Altman T.J.;

RT "Leptin receptor gene variation and obesity: lack of association in a
 RT white British male population."
 RL Hum. Mol. Genet. 6:869-876 (1997).
 RN [1]
 RP FUNCTION: Receptor for obesity factor (leptin). Involved in the
 CC regulation of fat metabolism and in a hematopoietic pathway
 CC required for normal lymphopoiesis. May play a role in
 CC reproduction.
 CC [1]
 CC FUNCTION: The short form (isoform A) may act to transport leptin
 CC to the cerebrospinal fluid (by similarity).
 CC [1]
 CC SUBCELLULAR LOCATION: Type I membrane protein. Except for form E
 CC which could be soluble.
 CC [1]
 CC ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=5;
 CC Name=B; Synonyms=13.2, OBBD;
 CC Name=A; Synonyms=6.4, HUB219.3;
 CC Name=C; Synonyms=12.1, OBRA;
 CC Name=D; Synonyms=HUB219.2;
 CC Name=E; Synonyms=VSP_001693, VSP_001694;
 CC Name=F;
 CC IsoId=P48357-5; Sequence=VSP_001688;
 CC [1]
 CC TISSUE SPECIFICITY: Isoform A expressed in fetal liver and in
 CC hematopoietic tissues and choroid plexus. In adults highest
 CC expression in heart, liver, small intestine, prostate and ovary.
 CC low level in lung and kidney. Isoform B is highly expressed in
 CC hypothalamus.
 CC [1]
 CC DOMAIN: The cytoplasmic domain may be essential for intracellular
 CC signal transduction by activation of JAK tyrosine kinase and
 CC STATs.
 CC [1]
 CC PTM: Phosphorylated on two tyrosine residues. Tyr-966 may be the
 CC major site of phosphorylation. Phosphorylation on both sites is
 CC required for full activity (by similarity).
 CC [1]
 CC SIMILARITY: Belongs to the type I cytokine family of receptors.
 CC Subfamily 2.
 CC [1]
 CC SIMILARITY: Contains 3 fibronectin type III domains.
 CC [1]
 CC SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC [1]
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC [1]
 CC EMBL; U43168; AAA93015.1; -
 CC EMBL; U66495; AAB07495.1; -
 CC EMBL; U66496; AAB07496.1; -
 CC EMBL; U66497; AAB07497.1; -
 CC EMBL; U59263; AAB09673.1; -
 CC EMBL; U59248; AAB09673.1; JOINED.
 CC EMBL; U59249; AAB09673.1; JOINED.
 CC EMBL; U59250; AAB09673.1; JOINED.
 CC EMBL; U59252; AAB09673.1; JOINED.
 CC EMBL; U59253; AAB09673.1; JOINED.
 CC EMBL; U59254; AAB09673.1; JOINED.
 CC EMBL; U59255; AAB09673.1; JOINED.
 CC EMBL; U59256; AAB09673.1; JOINED.
 CC EMBL; U59257; AAB09673.1; JOINED.
 CC EMBL; U59258; AAB09673.1; JOINED.
 CC EMBL; U59259; AAB09673.1; JOINED.
 CC EMBL; U59260; AAB09673.1; JOINED.
 CC EMBL; U59261; AAB09673.1; JOINED.
 CC EMBL; U59262; AAB09673.1; JOINED.
 CC EMBL; U50748; AAC23650.1; -
 CC EMBL; U52912; AAC50509.1; -
 CC EMBL; U52913; AAC50510.1; -
 CC EMBL; U52914; AAC50511.1; -
 CC HSSP; P16471; IBP3.
 CC GeneW; HGNC:6554; LEPR.
 CC MIW; 601007; -

```

DR GO: GO:0004888; P:transmembrane receptor activity; TAS.
DR GO: GO:0007166; P:cell surface receptor linked signal transdu. . .; TAS.
DR GO: GO:0007275; P:development; TAS.
DR GO: GO:0006112; P:energy reserve metabolism; TAS.
DR InterPro: IPR002996; CRIA.
DR InterPro: IPR008957; FN_III-like.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003529; Hemicoptcn_L_F2.
DR InterPro: IPR007110; Ig-like.
DR Pfam: PF00041; fn3_2.
DR SMART: SM00060; FN3_4.
DR PROSITE: PS01353; HEMATOPO_REC_L_F2_1.
DR PROSITE: PS50835; IG LIKE_1.
DR Oeasly: Receptor; Transmembrane; Glycoprotein; Signal;
KW Immunoglobulin domain; Repeat; Alternative splicing; Phosphorylation;
KW Polymorphism.
KM SIGNAL.
FT CHAIN 1 21
FT DOMAIN 22 1165
FT TRANSMEM 22 839
FT DOMAIN 840 862
FT DOMAIN 863 1165
CYTOPLASMIC (POTENTIAL).

Query Match 1.1%; Score 7; DB 1; Length 1165;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPV 377
DB 169 LEDSPV 175

RESULT 42
DHX8_ARATH STANDARD; PRT; 1168 AA.
ID DHX8; Q9LRV0;
AC Q3853; Q9LRV0;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative pre-mRNA splicing factor ATP-dependent RNA helicase.
GN ATG36560 OR MEI6.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RX SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=97086699; PubMed=8932388;
RA Quigley F., Dao P., Cotter A., Mache R.;
RT "Sequence analysis of an 81 kb contig from Arabidopsis thaliana
RT chromosome III."
RL Nucleic Acids Res. 24:4313-4318(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:131-135(2000).
CC -1- FUNCTION: Could act late in the splicing of pre-mRNA and mediate
CC the release of the spliced mRNA from spliceosomes (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DEAD
CC SUBFAMILY. DDX8/PRP22 ORTHOLOG.
CC -1- SIMILARITY: Contains 1 SI motif domain.
CC -1- CAUTION: Ref.1 sequences differ from that shown due to a
CC frameshift in position 57.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL: X98130; CAA66825.1; ALT FRAME.
DR EMBL: X97970; CAA66613.1; ALT FRAME.
DR EMBL: AB028611; BAB01838.1; -.
DR HSP: P05055; ISRO.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002464; DEAD box.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR007502; Helicase_dom.
DR InterPro: IPR008984; Nucleic_acid_OB.
DR InterPro: IPR003029; S1.
DR Pfam: PF04408; HA2_1.
DR Pfam: PF00271; helicase_C_1.
DR Pfam: PF00575; S1_1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC; 1.
DR SMART: SM00316; S1_1.
DR PROSITE: PS00690; DEAD_ATP_HELICASE_1.
DR PROSITE: PS50126; S1_1.
KW Hypothetical protein; mRNA processing; mRNA splicing; Helicase;
KW ATP-binding; Nuclear protein.
KM ATP-BINDING; Nuclear protein.
FT DOMAIN 214 283
FT NE_BIND 538 545
FT SITE 635 638
FT SITE 777 780
FT DOMAIN 777 780
FT POLY-PRO.
SQ SEQUENCE 1168 AA; 134156 MW; B3632DEA4A7A7690C CRC64;

```

```

Query Match 1.1%; Score 7; DB 1; Length 1168;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 KIMKRL 58
DB 653 KIMKRL 659

RESULT 43
AIP1_MOUSE STANDARD; PRT; 1275 AA.
ID AIP1_MOUSE
AC Q9W0L1; Q8BYT1; Q8CAB5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Atrophin-1 interacting protein 1 (Membrane associated guanylate kinase
DE inverted-2) (MAGI-2) (Activin receptor interacting protein 1)
DE (Acvripl).
GN AIP1 OR MAGI2 OR AIP1 OR ACVRINP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), INTERACTION WITH MADH2, MADH3
RP AND ACV22, AND IDENTIFICATION IN A COMPLEX WITH ACV22; ACV1B AND
RP MADH3.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE=20148748; PubMed=10681527;
RA Shoji H., Tsuchida K., Kishi H., Yamakawa N., Matsuzaki T., Liu Z.,
RA Nakamura T., Sugino H.;
RT "Identification and characterization of a PDZ protein that interacts
RT with activin types II receptors."
RL J. Biol. Chem. 275:5485-5492(2000).
RN [2]
RP SEQUENCE OF 159-1112 FROM N.A. (ISOFORM 3).
RC STRAIN=57BL/6J; TISSUE=Hypothalamus, and Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Salto R., Suzuki H., Yamakawa N., Kiyosawa H.,

```

RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldairelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schiraldi L.M., Kanapin A., Matsuda H., Batalov S., Betzel K.W.,
RA Blake J.A., Bratt D., Bruscia V., Chochua C., Corbani L.B., Cousins S.,
RA Datta E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Guetlich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochik I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Meglert D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagasima T., Numa K., Okido T., Pavan W.J., Perisa G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verrario R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilmig L.G., Wyszynski-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arkawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterson R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
CC -1- FUNCTION: Seems to act as scaffold molecule at synaptic junctions
CC by assembling neurotransmitter receptors and cell adhesion
CC proteins. May play a role in regulating activin-mediated signaling
CC in neuronal cells. Enhances the ability of p15N to suppress AKT1
CC activation (By similarity).
CC -1- SUBUNIT: Interacts via its second PDZ domain with PTEN unphosphorylated C-
CC terminus (By similarity). Interacts through its guanylate kinase
CC domain with DGAP1 (By similarity). Interacts through the PDZ
CC domains with GRIN2A, GRIN2B and NLGN1 (By similarity). Interacts
CC with CTNND2, CTNND1 and MAGUIN-1 (By similarity). Interacts with
CC ACVR2, MAD2 and MAD3. Part of a complex consisting of Alpi,
CC ACVR2, ACVR1B and MAD3.
CC -1- SUBCELLULAR LOCATION: Membrane-associated in synaptosomes (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1; Synonyms=long;
CC IsoId=Q9WV01-1; Sequence=Displayed;
CC Name=2; Synonyms=short;
CC IsoId=Q9WV01-2; Sequence=VSP_008436;
CC Note=Major;
CC Name=3;
CC IsoId=Q9WV01-3; Sequence=VSP_008437;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Specifically expressed in brain.
CC -1- SIMILARITY: Belongs to the MAGUK family.
CC -1- SIMILARITY: Contains 1 guanylate kinase-like domain.
CC -1- SIMILARITY: Contains 6 PDZ/DHR domains.
CC -1- CAUTION: Ref.2 (BAC29987) sequence differs from that shown due to
CC a frameshift in position 102.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL; AB029485; BAA82294.1; -;
CC EMBL; AK039336; BAC30321.1; ALT_INT.
CC EMBL; AK038407; BAC29987.1; ALT_FRAME.
CC PIR; P70546; P70698.
CC HSP; P29476; IQAV.

DR MCD; MG1.1354953; Acyrinpl.
DR GO; GO:0004871; F-signal transducer activity; IPI.
DR GO; GO:0007165; P-signal transduction; IPI.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001202; WW_Rsp5_WMP.
DR Pfam; PF00595; PDZ; 5.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00228; PDZ; 5.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS00856; GUANYLATE KINASE 1; 1.
DR PROSITE; PS50052; GUANYLATE KINASE 2; 1.
DR PROSITE; PS50106; PDZ; 6.
DR PROSITE; PS01159; WW_DOMAIN 1; 2.
DR PROSITE; PS50020; WW_DOMAIN 2; 2.
KW Repeat; Alternative splicing.
FT DOMAIN 17 101 PDZ 1.
FT DOMAIN 109 285 GUANYLATE KINASE.
FT DOMAIN 301 334 WW 1.
FT DOMAIN 347 380 WW 2.
FT DOMAIN 425 509 PDZ 2.
FT DOMAIN 604 682 PDZ 3.
FT DOMAIN 777 859 PDZ 4.
FT DOMAIN 919 1009 PDZ 5.
FT DOMAIN 1139 1221 PDZ 6.
FT VARSPPLIC 1 163 Missing (in isoform 2).
FT VARSPPLIC 1229 1275 /FTId=VSP_008436.
FT VARSPPLIC 1275 1275 WPSLSLWCMKSDHGSEFYFLALHPKDTNPFGVLPPLP
FT POACRK -> AFHGFHLCSAFSVF (in isoform 3).
FT FTId=VSP_008437.
SQ SEQUENCE 1275 AA; 140918 MW; F17DC52517806354 CRC64;
Query Match 1.1%; Score 7; DB 1; Length 1275;
Beet Local Similarity 100.0%; Pred. No. 2,4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 ALLESGT 23
Db 171 ALLESGT 177
RESULT 44
ALPI_RAT
ID ALPI_RAT STANDARD; PRT; 1277 AA.
AC 088382; Q9R271;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Atrophin-1 interacting protein 1 (Membrane associated guanylate kinase
DE inverted-2) (MAGI-2) (Synaptic scaffolding molecule) (S-SCAM).
GN ALPI OR MAGI2 OR SSCAM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, SUBCELLULAR LOCATION, TISSUE
RP SPECIFICITY, AND INTERACTION WITH DGAP1, NLGN1 AND GRIN2A.
RX MEDLINE=96361985; PubMed=9694864;
RA Hirao K., Hata Y., Ide N., Takeuchi M., Irie M., Yao I., Deguchi M.,
RA Toyoda A., Sudhof T.C., Takai Y.,
RA "A novel multiple PDZ domain-containing molecule interacting with
RA N-methyl-D-aspartate receptors and neuronal cell adhesion proteins.";
RL J. Biol. Chem. 273:21105-21110(1998).
[2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND ALTERNATIVE SPLICING (ISOFORM 3).
RX MEDLINE=20112872; PubMed=10644767;
RA Hirao K., Hata Y., Yao I., Deguchi M., Kawabe H., Mizoguchi A.,
RA Takai Y.,
RT "Three isoforms of synaptic scaffolding molecule and their
RT characterization: Multimerization between the isoforms and their
RT interaction with N-methyl-D-aspartate receptors and SAP90/PSD-95-
RT associated protein.";
RL J. Biol. Chem. 275:2966-2972(2000).

[3]
 RN INTERACTION WITH MAGUIN-1.
 RX MEDLINE=9922314; PubMed=10207009;
 RA Yao I., Hata Y., Ide N., Hirao K., Deguchi M., Nishio H.,
 RA Mizoguchi A., Takai Y.,
 RT "MAGUIN, a novel neuronal membrane-associated guanylate
 RT kinase-interacting protein."
 RL J. Biol. Chem. 274:11889-11896(1999).
 RN [4]
 RN INTERACTION WITH CTNND2.
 RX MEDLINE=99182311; PubMed=10080919;
 RA Ide N., Hata Y., Deguchi M., Hirao K., Yao I., Takai Y.,
 RT "Interaction of S-SCAM with neural plakophilin-related
 RT Armadillo-repeat protein/delta-catenin."
 RL Biochem. Biophys. Res. Commun. 256:456-461(1999).
 RN [5]
 RN INTERACTION WITH GRID2.
 RX MEDLINE=22477859; PubMed=12589829;
 RA Yip C.C., Muto Y., Kishida H., Hashikawa T., Yano R.,
 RT "PKC regulates the delta2 glutamate receptor interaction with
 RT S-SCAM/MAGI-2 protein."
 RL Biochem. Biophys. Res. Commun. 301:1122-1128(2003).
 CC -1- FUNCTION: Seems to act as scaffold molecule at synaptic junctions
 CC by assembling neurotransmitter receptors and cell adhesion
 CC proteins. May play a role in regulating activin-mediated signaling
 CC in neuronal cells. Enhances the ability of PTEN to suppress AKT1
 CC activation (By similarity).
 CC -1- SUBUNIT: Interacts through guanylate kinase domain its with
 CC DLGAP1. Interacts through the PDZ domains with GRIIN2A, GRID2 and
 CC NNG1. Interacts with CTNND2 and MAGUIN-1. Interacts with CTNNB1,
 CC ACVR2, MADH2 and MADH3 (By similarity). Part of a complex
 CC consisting of AIP1, ACVR2, ACVR1B and MADH3 (By similarity).
 CC Interacts via its WW domain with DRPLA (By similarity). Interacts
 CC via its second PDZ domain with PTEN unphosphorylated C-terminus
 CC (By similarity).
 CC -1- SUBCELLULAR LOCATION: Membrane-associated in synapses.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Name=1; Synonym=alpha;
 CC IsoId=088382-1; Sequence=Displayed;
 CC Name=2; Synonym=beta;
 CC IsoId=088382-2; Sequence=VSP_008438;
 CC Name=3; Synonym=gamma;
 CC IsoId=088382-3; Sequence=VSP_008439;
 CC -1- TISSUE SPECIFICITY: Exclusively expressed in brain.
 CC -1- SIMILARITY: Belongs to the MAGUK family.
 CC -1- SIMILARITY: Contains 1 guanylate kinase-like domain.
 CC -1- SIMILARITY: Contains 6 PDZ/DHR domains.
 CC -1- SIMILARITY: Contains 2 WW domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

DR PROSITE; PSS0106; PDZ; 6.
 DR PROSITE; PSS01159; WW DOMAIN 1; 2.
 DR PROSITE; PSS0020; WW DOMAIN 2; 2.
 KM Repeat; Alternative Splicing; PDZ.
 FT DOMAIN 109 101
 FT DOMAIN 285 285
 FT DOMAIN 302 335
 FT DOMAIN 348 381
 FT DOMAIN 426 510
 FT DOMAIN 605 683
 FT DOMAIN 778 860
 FT DOMAIN 920 1010
 FT DOMAIN 1141 1223
 FT VARSPPLIC 1 163
 FT VARSPPLIC 1 223
 FT VARSPPLIC 645 645
 FT CONFLICT 1277 AA; 141071 MW; E1A435FF3549DF9 CRC64;
 SQ SEQUENCE 1277 AA; 141071 MW; E1A435FF3549DF9 CRC64;
 Query Match 1.1%; Score 7; DB 1; Length 1277;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 17 ALLESGT 23
 Db 171 ALLESGT 177
 RESULT 45
 YAB2_SCHPO STANDARD; PRT; 1318 AA.
 ID YAB2_SCHPO
 AC 0098064;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C2G11.02 in chromosome I.
 GN SPAC2G11.02.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Baaham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagals K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkstree E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Welljens I., Vansteelandt E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleux V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet L., Gallard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RT "The genome sequence of Schizosaccharomyces pombe".
 RL Nature 415:871-880(2002).

CC colobomas of the lower eyelids; (3) hypoplasia of the mandible and
 CC zygomatic complex; (4) cleft palate.
 CC - - SIMILARITY: Contains 1 Lish domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U40847; AAC50903.1; -
 DR EMBL; U76366; AAC51181.1; -
 DR EMBL; U84664; AAC51185.1; -
 DR EMBL; U84640; AAC51185.1; JOINED.
 DR EMBL; U84641; AAC51185.1; JOINED.
 DR EMBL; U84642; AAC51185.1; JOINED.
 DR EMBL; U84643; AAC51185.1; JOINED.
 DR EMBL; U84644; AAC51185.1; JOINED.
 DR EMBL; U84645; AAC51185.1; JOINED.
 DR EMBL; U84646; AAC51185.1; JOINED.
 DR EMBL; U84647; AAC51185.1; JOINED.
 DR EMBL; U84648; AAC51185.1; JOINED.
 DR EMBL; U84649; AAC51185.1; JOINED.
 DR EMBL; U84650; AAC51185.1; JOINED.
 DR EMBL; U84651; AAC51185.1; JOINED.
 DR EMBL; U84652; AAC51185.1; JOINED.
 DR EMBL; U84653; AAC51185.1; JOINED.
 DR EMBL; U84654; AAC51185.1; JOINED.
 DR EMBL; U84655; AAC51185.1; JOINED.
 DR EMBL; U84656; AAC51185.1; JOINED.
 DR EMBL; U84657; AAC51185.1; JOINED.
 DR EMBL; U84658; AAC51185.1; JOINED.
 DR EMBL; U84659; AAC51185.1; JOINED.
 DR EMBL; U84660; AAC51185.1; JOINED.
 DR EMBL; U84661; AAC51185.1; JOINED.
 DR EMBL; U84662; AAC51185.1; JOINED.
 DR EMBL; U84663; AAC51185.1; JOINED.
 DR EMBL; U79659; AAB40722.1; -
 DR EMBL; U79645; AAB40722.1; JOINED.
 DR EMBL; U79646; AAB40722.1; JOINED.
 DR EMBL; U79647; AAB40722.1; JOINED.
 DR EMBL; U79648; AAB40722.1; JOINED.
 DR EMBL; U79649; AAB40722.1; JOINED.
 DR EMBL; U79650; AAB40722.1; JOINED.
 DR EMBL; U79651; AAB40722.1; JOINED.
 DR EMBL; U79652; AAB40722.1; JOINED.
 DR EMBL; U79653; AAB40722.1; JOINED.
 DR EMBL; U79654; AAB40722.1; JOINED.
 DR EMBL; U79655; AAB40722.1; JOINED.
 DR EMBL; U79656; AAB40722.1; JOINED.
 DR EMBL; U79657; AAB40722.1; JOINED.
 DR EMBL; U79658; AAB40722.1; JOINED.
 DR Genew; HGNC:11654; TCOR1.
 DR MIM; 606847; -
 DR MIM; 154500; -
 DR GO; GO:0005730; C:nucleolus; TAS.
 DR GO; GO:0005215; P:transporter activity; TAS.
 DR GO; GO:0001501; P:skeletal development; TAS.
 DR InterPro: IPR006594; Lish.
 DR InterPro: IPR003993; trectacle.
 DR Pfam; PF03946; trectacle.3.
 DR PRINTS; PRO1503; TREACLE.
 DR SMART; SM00667; Lish; 1.
 DR PROSITE; PS50896; LISH; 1.
 KM Disease mutation; Polymorphism.
 FT DOMAIN 6 38 LISH.
 FT DOMAIN 89 97 POLY-GU.
 FT DOMAIN 204 207 POLY-SER.
 FT DOMAIN 616 619 POLY-SER.
 FT DOMAIN 919 924 POLY-SER.
 FT DOMAIN 1285 1289 POLY-LYS.

FT DOMAIN 1375 1386 POLY-LYS.
 FT DOMAIN 1398 1405 POLY-LYS.
 FT VARIANT 53 53 W -> R (in TCS).
 FT VARIANT 439 439 /FTID=VAR_005630.
 FT VARIANT 439 439 P -> L.
 FT VARIANT 810 810 /FTID=VAR_005631.
 FT VARIANT 810 810 A -> V.
 FT VARIANT 1313 1313 /FTID=VAR_005632.
 FT VARIANT 1313 1313 A -> V (in absNP:15251).
 FT VARIANT 1355 1355 /FTID=VAR_005633.
 FT VARIANT 1355 1355 D -> G.
 FT CONFLICT 1312 1312 /FTID=VAR_005634.
 FT CONFLICT 1411 AA; 144312 MM; 3880203D985C2699 CRC64; K -> Q (IN REF. 2).
 SQ SEQUENCE 1411 AA; 144312 MM; 3880203D985C2699 CRC64;
 Query Match. 1.1%; Score 7; DB 1; Length 1411;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 357 PASVPSV 363
 Db 767 PASVPSV 773
 RESULT 48
 ID AIP1_HUMAN STANDARD; PRT; 1455 AA.
 AC Q86UJ8; O60434; O60510; Q86UJ7; Q9UDQ5; Q9UDU1;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Atrophin-1 interacting protein 1 (Atrophin-1 interacting protein A)
 DE (Membrane associated guanylate kinase inverted-2) (MAGI-2).
 OS Homo sapiens (Human).
 GN AIP1 OR MAGI2 OR KIAA0705.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, AND INTERACTION
 RP WITH DRPLA.
 RC TISSUE=Brain;
 RX MEDLINE=98313405; PubMed=9647693;
 RA Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan X., Kushi J.,
 RA Kaminsky Z., Kleiderlein J.U., Sharp A.H., Rose C.A.;
 RA "Atrophin-1, the DRPLA gene product, interacts with two families of WW
 RA domain-containing proteins.";
 RL Mol. Cell. Neurosci. 11:149-160(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX MEDLINE=98403880; PubMed=9734811;
 RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
 RA Kottani H., Nomura N., Ohara O.;
 RA "Prediction of the coding sequences of unidentified human genes. X.
 RA The complete sequences of 100 new cDNA clones from brain which can
 RA code for large proteins in vitro.";
 RL DNA Res. 5:169-176(1998).
 RN [3]
 RP SEQUENCE OF 1-100; 141-179; 350-682; 684-1175 AND 1237-1455 FROM N.A.
 RA Courtney L., Harrison M., Lennox S., Bourne S., Ozeresky P., Carter T.,
 RA Tin-Mollam A., Hawkins M., Kepler D., Sulston J.E., Eldred J.,
 RA Kozlowicz A., Bemis G., Langston Y., Mead K., Haekenson W., Meyer R.,
 RA Shih K., Waterston R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP FUNCTION, SUBCELLULAR LOCATION, AND INTERACTION WITH PTEN.
 RX MEDLINE=20226087; PubMed=10760291;
 RA Wu X., Hepler K., Caecelino-Prabhu S., Do D., Kaye M.B., Yuan X.-J.,
 RA Wood J., Ross C., Sawyers C.L., Wang Y.E.;
 RA "Evidence for regulation of the PTEN tumor suppressor by a
 RA membrane-localized multi-PDZ domain containing scaffold protein
 RA MAGI-2.";

Proc. Natl. Acad. Sci. U.S.A. 97:4233-4238(2000).
 [5]
 INTERACTION WITH PTEN.
 MEDLINE=21623681; PubMed=11707428;
 RA Valquez F., Grossman S.R., Takahashi Y., Rokas M.V., Nakamura N.,
 RA Sellers W.R.;
 RT "Phosphorylation of the PTEN tail acts as an inhibitory switch by
 preventing its recruitment into a protein complex.";
 RL J. Biol. Chem. 276:48627-48630(2001).
 CC -1- FUNCTION: Seems to act as scaffold molecule at synaptic junctions
 by assembling neurotransmitter receptors and cell adhesion
 proteins. May play a role in regulating activin-mediated signaling
 in neuronal cells. Enhances the ability of PTEN to suppress Akt1
 activation.
 CC -1- SUBUNIT: Interacts via its WW domains with DRPLA. Interacts via
 its second PDZ domain with PTEN unphosphorylated C-terminus.
 CC Interacts through its guanylate kinase domain with DLGAP1 (By
 similarity). Interacts through the PDZ domains with GRIN2A, GRID2
 and NLGN1 (By similarity). Interacts with CTNND2, CTNNB1,
 MAGUIN-1, ACVR2, MADH2 and MADH3 (By similarity). Part of a
 complex consisting of AIPL1, ACVR2, ACVR1B and MADH3 (By
 similarity).
 CC -1- SUBCELLULAR LOCATION: Membrane-associated in synaptosomes (By
 similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q86UL8-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q86UL8-2; Sequence=VSP_008435;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Specifically expressed in brain.
 CC -1- SIMILARITY: Belongs to the MAGUK family.
 CC -1- SIMILARITY: Contains 1 guanylate kinase-like domain.
 CC -1- SIMILARITY: Contains 6 PDZ/DHR domains.
 CC -1- SIMILARITY: Contains 2 WW domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF038563; AAC05370.1; -;
 DR EMBL; AB014605; BAA31680.2; ALT_INIT.
 DR EMBL; AC004808; AAC23438.1; -;
 DR EMBL; AC004945; AAC61488.1; -;
 DR EMBL; AC004990; AAC79151.1; -;
 DR EMBL; AC005246; AAC25530.1; -;
 DR EMBL; AC006043; AAD15413.2; -;
 DR EMBL; AC073200; AAP22360.1; -;
 DR EMBL; AC007237; AAP21886.1; -;
 DR HSSP; P29476; IQAV.
 DR MIM; 606382; -;
 DR InterPro; IPR01202; WW_Rsp5_WWP.
 DR Pfam; PF00625; Guanylate_kin; 1.
 DR Pfam; PF00397; WW; 2.
 DR SMART; SM00456; WW; 1.
 DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
 DR PROSITE; PS0052; GUANYLATE_KINASE_2; 1.
 DR PROSITE; PS0106; PDZ; 6.
 DR PROSITE; PS0159; WW_DOMAIN_1; 2.
 DR PROSITE; PS0020; WW_DOMAIN_2; 2.
 KW Repeat; Alternative splicing.
 FT DOMAIN 17 101 PDZ 1.
 FT DOMAIN 109 285 GUANYLATE_KINASE.
 FT DOMAIN 302 335 WW 1.
 FT DOMAIN 348 381 WW 2.
 FT DOMAIN 426 510 PDZ 2.
 FT DOMAIN 605 683 PDZ 3.

FT	DOMAIN	778	860	PDZ 4.
FT	DOMAIN	920	1010	PDZ 5.
FT	DOMAIN	1015	1118	PRO-RICH.
FT	DOMAIN	1147	1229	PDZ 6.
FT	DOMAIN	1340	1430	ALA-RICH.
FT	VARSPLIC	757	771	OOVPRTPSPRMSSG -> R (in isoform 2).
FT	CONFLICT	1234	1234	/FTId=VSP_008435.
FT	CONFLICT	1250	1250	Q -> E (IN REF. 2).
FT	CONFLICT	1291	1291	G -> C (IN REF. 1).
FT	CONFLICT	1383	1383	E -> K (IN REF. 1).
FT	CONFLICT	1389	1394	P -> L (IN REF. 1).
FT	CONFLICT	1401	1401	FAGPGG -> SADPAD (IN REF. 1).
FT	CONFLICT	1411	1411	E -> A (IN REF. 1).
FT	CONFLICT	1415	1415	G -> A (IN REF. 1).
FT	CONFLICT	1420	1420	PG -> SV (IN REF. 1).
FT	CONFLICT	1423	1423	G -> A (IN REF. 1).
FT	CONFLICT	1426	1426	P -> R (IN REF. 1).
FT	CONFLICT	1429	1429	K -> R (IN REF. 1).
FT	CONFLICT	1437	1437	V -> G (IN REF. 1).
FT	CONFLICT	1437	1437	P -> R (IN REF. 1).
SQ	SEQUENCE	1455 AA;	158752 MW;	3AF8A14D69738281 CRC64;

Query Match 1.1%; Score 7; DB 1; Length 1455;
 Best Local Similarity 100.0%; Pred. No. 2,7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	17	ALLBSGT	23
DB	171	ALLBSGT	177

RESULT 49
 MYHB_HUMAN
 ID MYHB_HUMAN STANDARD; PRT; 1972 AA.
 AC P35749; O00396; O94944; P78422;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Myosin heavy chain, smooth muscle isoform (SMHHC).
 GN MYH1 OR KIA00866.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99425270; PubMed=10493829;
 RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush P., Brandon R.,
 RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
 RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
 RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
 RT "Genome duplications and other features in 12 Mb of DNA sequence from
 human chromosome 16p and 16q.";
 RL Genomics 60:295-308(1999).
 RN [2]
 RP SEQUENCE FROM N.A., AND REVISIONS TO 1263-1266.
 RC TISSUE=Brain;
 RA Nagase T., Kikuno R., Yamakawa H., Ohara O.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-1266 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99156230; PubMed=10048485;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XII.
 The complete sequences of 100 new cDNA clones from brain which code
 for large proteins in vitro.";
 RL DNA Res. 5:355-364(1998).
 RN [4]
 RP SEQUENCE OF 885-1972 FROM N.A.
 RX MEDLINE=9326189; PubMed=7684189;
 RA Matsuo R., Yoshida M.C., Furutani Y., Imamura S., Kanda N.,

RA Yanagisawa M., Masaki T., Takao A.;
 RT "human smooch muscle myosin heavy chain gene mapped to chromosomal
 RL region 16q12."; Am. J. Med. Genet. 46:61-67(1993).
 RN [5]
 RP SEQUENCE OF 1093-1972 FROM N.A.
 RC TISSUE=Hipocampus;
 RA Okajima K.;
 RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Muscle contraction.
 CC -1- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2
 CC heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)
 CC and 2 regulatory light chain subunits (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- TISSUE SPECIFICITY: Smooch muscle; expressed in the umbilical
 CC artery, bladder, esophagus and trachea.
 CC -1- DOMAIN: The rodlike tail sequence is highly repetitive, showing
 CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
 CC characteristic for alpha-helical coiled coils.
 CC -1- DISEASE: A chromosomal rearrangement, known as pericentric
 CC inversion inv(16)(p13;q22), produces a fusion protein that
 CC consists of the 165 N-terminal residues of CBF-beta (PEPB2) with
 CC the tail region of MYH11. This rearrangement is associated with
 CC acute myeloid leukemia of M4EO subtype.
 CC -1- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light
 CC meromyosin (LMW) and 1 heavy meromyosin (HMM). It can later be
 CC split further into 2 globular subfragments (S1) and 1 rod-shaped
 CC subfragment (S2).
 CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.
 CC -1- SIMILARITY: Contains 1 IQ domain.
 CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW=http://www.infobio.gen.fr/services/chromancer/genes/MYH11ID43.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF001548; AAC31665.1; -;
 CC EMBL; U91323; AAC35212.1; -;
 CC EMBL; AB020673; BA47489.2; ALT. INT.
 CC EMBL; D10667; -; NOT_ANNOTATED_CDS.
 CC EMBL; X69292; CAA49154.1; -;
 CC HSSP; P10587; IBR2.
 CC Genew; HGNC:7569; MYH11.
 CC MIM; 160745; -;
 CC DR GO; GO:0005859; C:muscle myosin; TAS.
 CC DR InterPro; IPR000048; IQ_region.
 CC DR InterPro; IPR001609; Myosin_head.
 CC DR InterPro; IPR004009; Myosin_N.
 CC DR InterPro; IPR002928; Myosin_tail.
 CC DR InterPro; IPR002017; Spectrin.
 CC Pfam; PF00612; IQ; 1.
 CC Pfam; PF00063; myosin_head; 1.
 CC Pfam; PF02736; Myosin_N; 1.
 CC Pfam; PF01576; Myosin_tail; 1.
 CC PRINTS; PR00193; MYOSINHEAVY.
 CC Prodom; PD000355; myosin_head; 1.
 CC SMART; SM00015; IQ; 1.
 CC SMART; SM00242; MYSC; 1.
 CC PROSITE; PS50096; IQ; 1.
 CC DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 CC KAModulin-binding; ATP-binding; Methylation; Multigene family;
 CC KM Proto-oncogene; Chromosomal translocation.
 CC FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
 CC FT DOMAIN 786 815 IQ.
 CC FT DOMAIN 844 1934 COILED COIL (POTENTIAL).
 CC FT DOMAIN 1935 1972 CARBOXYL-TERMINAL.
 CC NP BIND 178 185 ATP (POTENTIAL).
 CC FT DOMAIN 661 683 ACTIN-BINDING (BY SIMILARITY).

FT DOMAIN 762 776 ACTIN-BINDING (BY SIMILARITY).
 FT MOD RES 129 129 METHYLATION (TRI-) (POTENTIAL).
 FT CONFLICT 887 889 EER -> MSE (IN REF. 4).
 FT CONFLICT 1558 1558 T -> S (IN REF. 4).
 FT CONFLICT 1610 1611 KQ -> NE (IN REF. 4).
 FT CONFLICT 1786 1786 A -> S (IN REF. 5).
 FT CONFLICT 1958 1958 T -> L (IN REF. 4).
 SQ SEQUENCE 1972 AA; 227338 MW; 67655B2AEC81277 CRC64;
 Query Match 1.1%; Score 7; DB 1; Length 1972;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 94 DVASLSS 100
 Db 1309 DVASLSS 1315
 RESULT 50
 TEGU_HSV6G
 ID TEGU_HSV6G STANDARD; PRT; 2077 AA.
 AC P30002;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Large tegument protein.
 OS U31.
 GN Human herpesvirus (type 6 / strain GS) (HHV6).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Roseolovirus.
 OX NCBI_TaxID=10369;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91374623; PubMed=1654455;
 RA Josephs S.F., Ablashi D.V., Salahuddin S.Z., Jagodzinski L.L.,
 RA Wong-Staal F., Gallo R.C.;
 RT "Identification of the human herpesvirus 6 glycoprotein H and
 RT putative large tegument protein genes";
 RL J. VIROL. 65:5597-5604(1991).
 CC -1- FUNCTION: Tegument protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
 CC BHV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S57540; AAB19786.1; -;
 CC DR PIR; H40511; W2BE24.
 CC DR InterPro; IPR006928; Herpes teg N.
 CC DR Pfam; PF04843; Herpes teg N; 1.
 CC SQ SEQUENCE 2077 AA; 23909 MW; 28E297FDC2FAD2AE CRC64;
 Query Match 1.1%; Score 7; DB 1; Length 2077;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 601 OKRSDRA 607
 Db 691 OKRSDRA 697

Search completed: June 21, 2004, 13:27:27
 Job time : 21 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:23:28 ; Search time 45 Seconds
(without alignments)
4662.658 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665

Sequence: 1 MAHEMIGTQIVTERLVALE.....LGVKVSQSPSSGSMETIEVS 665

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 172362

Minimum DB seq length: 473

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database :

1: SPREMBL 25:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	71.0	665	4 Q96N49	Q96N49 homo sapien
2	371	55.8	616	4 Q8N571	Q8N571 homo sapien
3	304	45.7	662	4 Q8IVT8	Q8IVT8 mus sapien
4	67	10.1	660	11 Q920R2	Q920R2 mus musculu
5	64	9.6	677	11 Q99MG6	Q99MG6 mus musculu
6	48	7.2	622	11 Q99MG5	Q99MG5 mus musculu
7	20	3.0	625	4 Q86S88	Q86S88 homo sapien
8	20	3.0	665	11 Q7S29	Q7S29 mus musculu
9	13	2.0	476	5 Q9VH8	Q9VH8 drosophila
10	13	2.0	476	5 Q86122	Q86122 drosophila
11	9	1.4	606	5 Q8ST18	Q8ST18 caenorhabdi
12	9	1.4	657	5 Q8ST19	Q8ST19 caenorhabdi
13	8	1.2	544	16 Q82GP4	Q82GP4 streptomyce
14	8	1.2	559	16 Q8ABA4	Q8ABA4 bacteroides
15	8	1.2	694	16 Q8ZK16	Q8ZK16 salmonella
16	8	1.2	707	17 Q8U1R9	Q8U1R9 pyrococcus

17	8	1.2	727	16 Q8YP24	Q8YP24 anabaena sp
18	8	1.2	858	11 Q8K058	Q8K058 mus musculu
19	8	1.2	892	11 Q80T06	Q80T06 mus musculu
20	8	1.2	1001	11 Q8BMG7	Q8BMG7 mus musculu
21	8	1.2	1042	16 Q8Z0R6	Q8Z0R6 nitrosomona
22	8	1.2	1065	5 Q815U3	Q815U3 plasmodium
23	8	1.2	1194	10 Q9KFL1	Q9KFL1 arabidopsis
24	8	1.2	1308	5 Q01924	Q01924 caenorhabdi
25	8	1.2	1885	17 Q8RQD1	Q8RQD1 methanobarc
26	7	1.1	475	16 Q51322	Q51322 borrelia bu
27	7	1.1	478	5 Q9W131	Q9W131 drosophila
28	7	1.1	479	2 Q9S323	Q9S323 vibrio fisc
29	7	1.1	479	10 Q9LZG8	Q9LZG8 arabidopsis
30	7	1.1	480	2 Q56817	Q56817 photorhabdu
31	7	1.1	480	5 Q8SY11	Q8SY11 drosophila
32	7	1.1	483	4 Q9HEA7	Q9HEA7 homo sapien
33	7	1.1	487	10 Q9SXY1	Q9SXY1 arabidopsis
34	7	1.1	488	16 Q8ZJT2	Q8ZJT2 yersinia pe
35	7	1.1	490	16 Q81MG2	Q81MG2 bacillus an
36	7	1.1	492	2 Q9KXD8	Q9KXD8 escherichia
37	7	1.1	494	17 Q8ZYH4	Q8ZYH4 pyrobaculum
38	7	1.1	495	5 Q9W0U3	Q9W0U3 drosophila
39	7	1.1	499	5 Q76107	Q76107 culex quinq
40	7	1.1	502	5 Q9N3R4	Q9N3R4 caenorhabdi
41	7	1.1	505	12 Q806Y8	Q806Y8 cucumbe ye
42	7	1.1	508	16 Q7UH12	Q7UH12 rhodospirell
43	7	1.1	509	2 Q50188	Q50188 mycobacteri
44	7	1.1	509	16 Q9CDE5	Q9CDE5 mycobacteri
45	7	1.1	511	16 Q8YGD2	Q8YGD2 bruceella me
46	7	1.1	511	16 Q8G1J0	Q8G1J0 bruceella su
47	7	1.1	515	2 P77815	P77815 nocardioide
48	7	1.1	516	16 Q82NP2	Q82NP2 streptomyce
49	7	1.1	519	16 Q93N44	Q93N44 coxiella bu
50	7	1.1	523	4 Q8N3D9	Q8N3D9 homo sapien
51	7	1.1	523	11 Q9CU46	Q9CU46 mus musculu
52	7	1.1	526	10 Q86WS1	Q86WS1 homo sapien
53	7	1.1	526	4 Q9ZRH9	Q9ZRH9 oryza sativ
54	7	1.1	542	12 Q39481	Q39481 avian pneum
55	7	1.1	542	12 Q39479	Q39479 avian pneum
56	7	1.1	542	12 Q39478	Q39478 avian pneum
57	7	1.1	545	16 Q8A651	Q8A651 bacteroides
58	7	1.1	546	17 Q9Y9L7	Q9Y9L7 aeropyrum p
59	7	1.1	547	5 Q816X1	Q816X1 ancylostoma
60	7	1.1	548	16 Q9JWD7	Q9JWD7 neisseria m
61	7	1.1	551	16 Q8E4X9	Q8E4X9 streptococ
62	7	1.1	551	16 Q8D2B9	Q8D2B9 streptococ
63	7	1.1	555	2 Q8KX42	Q8KX42 streptomyce
64	7	1.1	555	16 Q8BSR7	Q8BSR7 lactobacill
65	7	1.1	564	16 Q9X144	Q9X144 thermocoga
66	7	1.1	574	11 Q8R1D2	Q8R1D2 mus musculu
67	7	1.1	576	10 Q9FL24	Q9FL24 arabidopsis
68	7	1.1	576	10 Q9SB22	Q9SB22 arabidopsis
69	7	1.1	578	16 Q9BDK0	Q9BDK0 rhizobium 1
70	7	1.1	580	16 Q82DF0	Q82DF0 streptomyce
71	7	1.1	596	4 Q9UPJ6	Q9UPJ6 homo sapien
72	7	1.1	601	16 Q86691	Q86691 streptomyce
73	7	1.1	605	6 Q7S590	Q7S590 sus scrofa
74	7	1.1	606	11 Q9ESS8	Q9ESS8 rattus norv
75	7	1.1	609	5 Q81VC3	Q81VC3 plasmodium

ALIGNMENTS

RESULT 1
Q96N49 PRELIMINARY; PRT; 665 AA.
AC Q96N49; DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein FLJ11411.
OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K.,
RA Iisaga T.,

RT "NDO human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK055973; BAB71060.1; -.

DR HSPG; Q16828; MKP.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.

DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.

DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.

DR PROSITE; PS50206; RHODANESE_3; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.

DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein; Hydrolase.
SQ SEQUENCE 665 AA; 73058 MW; 1EAEDDF08460DF CRC64;

Query Match 71.0%; Score 472; DB 4; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCCPDPFIPESHFLRPVNDSECEKILPWLKSVDFIEKAKSNGCVLHCLAGISRSAT 253
DB 194 TCCPDPFIPESHFLRPVNDSECEKILPWLKSVDFIEKAKSNGCVLHCLAGISRSAT 253
QY 254 IAIAYIMKRMDSLDEAYRFVKEKRTISPENFNLGQLDYEKKIKNOTGASGPKSKL 313
DB 254 IAIAYIMKRMDSLDEAYRFVKEKRTISPENFNLGQLDYEKKIKNOTGASGPKSKL 313
QY 314 LHLKNEPVPVAVSEGGQKSETPPLSPCADSATSEAAQGPVHPASVPSVPSVPSLLED 373
DB 314 LHLKNEPVPVAVSEGGQKSETPPLSPCADSATSEAAQGPVHPASVPSVPSVPSLLED 373
QY 374 SPLVQALSGHLSDADLEDSNKLKRSFLDIKSVYSASMAASLHGSSSEDALEYKPS 433
DB 374 SPLVQALSGHLSDADLEDSNKLKRSFLDIKSVYSASMAASLHGSSSEDALEYKPS 433
QY 434 TTLDTGNKLCQFSVQVSELTSEQTPETSPDKEASIPKKLOTARPSDSQSKLHVSRTSSG 493
DB 434 TTLDTGNKLCQFSVQVSELTSEQTPETSPDKEASIPKKLOTARPSDSQSKLHVSRTSSG 493
QY 494 TAOSSLSPHLRSGSVEDNHTSFLFGLSTSOQHLTKSAGLGLKGMHSDTLAPOTSTPSL 553
DB 494 TAOSSLSPHLRSGSVEDNHTSFLFGLSTSOQHLTKSAGLGLKGMHSDTLAPOTSTPSL 553
QY 554 TSSWYFATSSHPSASAIYGGASVAYSCSOLPTCGDQVYVRRRQKPSDRADSRSSW 613
DB 554 TSSWYFATSSHPSASAIYGGASVAYSCSOLPTCGDQVYVRRRQKPSDRADSRSSW 613
QY 614 HESPEPEKOPKRRSCOMEFGESIMSENRREELGKVGSGSFSGSMETIEVS 665
DB 614 HESPEPEKOPKRRSCOMEFGESIMSENRREELGKVGSGSFSGSMETIEVS 665

RESULT 2
ID Q8N5T1 PRELIMINARY; PRT; 616 AA.
AC Q8N5T1;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Lung;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031643; AAH31643.1; -.

DR GO; GO:0017017; F:MAP kinase phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.

DR InterPro; IPR000340; DS phosphatase.
DR InterPro; IPR000387; MAPK_phosph.
DR InterPro; IPR000387; TYR_phosphatase.

DR Pfam; PF00782; DSPC; 1.
DR PRINTS; PR01764; MAPKPHPTASE.
DR SMART; SM00195; DSPC; 1.

DR PROSITE; PS50056; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.

KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 616 AA; 67636 MW; 2CB0B14482F2AD72 CRC64;

Query Match 55.8%; Score 371; DB 4; Length 616;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 194 TCCPDPFIPESHFLRPVNDSECEKILPWLKSVDFIEKAKSNGCVLHCLAGISRSAT 253
DB 145 TCCPDPFIPESHFLRPVNDSECEKILPWLKSVDFIEKAKSNGCVLHCLAGISRSAT 204
QY 254 IAIAYIMKRMDSLDEAYRFVKEKRTISPENFNLGQLDYEKKIKNOTGASGPKSKL 313
DB 205 IAIAYIMKRMDSLDEAYRFVKEKRTISPENFNLGQLDYEKKIKNOTGASGPKSKL 264
QY 314 LHLKNEPVPVAVSEGGQKSETPPLSPCADSATSEAAQGPVHPASVPSVPSVPSLLED 373
DB 265 LHLKNEPVPVAVSEGGQKSETPPLSPCADSATSEAAQGPVHPASVPSVPSVPSLLED 324
QY 374 SPLVQALSGHLSDADLEDSNKLKRSFLDIKSVYSASMAASLHGSSSEDALEYKPS 433
DB 325 SPLVQALSGHLSDADLEDSNKLKRSFLDIKSVYSASMAASLHGSSSEDALEYKPS 384
QY 434 TTLDTGNKLCQFSVQVSELTSEQTPETSPDKEASIPKKLOTARPSDSQSKLHVSRTSSG 493
DB 385 TTLDTGNKLCQFSVQVSELTSEQTPETSPDKEASIPKKLOTARPSDSQSKLHVSRTSSG 444
QY 494 TAOSSLSPHLRSGSVEDNHTSFLFGLSTSOQHLTKSAGLGLKGMHSDTLAPOTSTPSL 553
DB 445 TAOSSLSPHLRSGSVEDNHTSFLFGLSTSOQHLTKSAGLGLKGMHSDTLAPOTSTPSL 504
QY 554 TSSWYFATSSHPSASAIYGGASVAYSCSOLPTCGDQVYVRRRQKPSDRADSRSSW 613
DB 505 TSSWYFATSSHPSASAIYGGASVAYSCSOLPTCGDQVYVRRRQKPSDRADSRSSW 564
QY 614 HESPEPEKOPKRRSCOMEFGESIMSENRREELGKVGSGSFSGSMETIEVS 665
DB 565 HESPEPEKOPKRRSCOMEFGESIMSENRREELGKVGSGSFSGSMETIEVS 616

RESULT 3
ID Q81VT8 PRELIMINARY; PRT; 662 AA.
ID Q81VT8

AC Q81V78; (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Similar to dual specificity phosphatase 16.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ductodenum;
 RA Strausberg R.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC042101; AAH42101.1; -
 DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR GO: GO:0006470; P:Protein amino acid dephosphorylation; IEA.
 DR InterPro: IPR000340; DS phosphatase.
 DR InterPro: IPR000343; MAPK phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSpC; 1.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKDPHTASE.
 DR SMART: SM00195; DSpC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PSS0206; RHODANSE_3; 1.
 DR PROSITE: PSS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PSS0056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PSS0054; TYR_PHOSPHATASE_DUAL; 1.
 DR HYDROLASE: Kinase.
 SQ SEQUENCE 662 AA; 72818 MW; 350534EF0652B98F CRC64;

Query Match 45.7%; Score 304; DB 4; Length 662;
 Best Local Similarity 100.0%; Pred. No. 3e-313;

Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 SVPSVOPSLDESPVLVQALSGLHLSADRLSDSNLKRSPFLDIKSVSYASMAASLHGFS 421
 DB 359 SVPSVOPSLDESPVLVQALSGLHLSADRLSDSNLKRSPFLDIKSVSYASMAASLHGFS 418
 QY 422 SSEDALERYKPTLLDGTNKLQFSPVQLSEQTPETSPDKEASIPKQLQTARPSDSQS 481
 DB 419 SSEDALERYKPTLLDGTNKLQFSPVQLSEQTPETSPDKEASIPKQLQTARPSDSQS 478
 QY 482 KRLHSVRTSSSGTARSLPLHRSQSVENNYHTSPFGLSTGQHLTKSAGLKGWHS 541
 DB 479 KRLHSVRTSSSGTARSLPLHRSQSVENNYHTSPFGLSTGQHLTKSAGLKGWHS 538
 QY 542 DILAPOTSTSLSSWFATSSHFYSASAIYGSASAYSCSOLPTGDOYYSVRRQ 601
 DB 539 DILAPOTSTSLSSWFATSSHFYSASAIYGSASAYSCSOLPTGDOYYSVRRQ 598
 QY 602 KPDRADSRSSWHEESPFKQFKRSCQMEFGESINSRRELGLGVSGSSFFSGSMEI 661
 DB 599 KPDRADSRSSWHEESPFKQFKRSCQMEFGESINSRRELGLGVSGSSFFSGSMEI 658
 QY 662 IEVS 665
 DB 659 IEVS 662

RESULT 4

Q920R2 PRELIMINARY; PRT; 660 AA.
 AC Q920R2; (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE MAP kinase phosphatase-7.
 GN DUSP16 OR 3830417MI7RIK OR MKP-7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cerebellum;
 RX MEDLINE=21486429; PubMed=11489891;
 RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
 RT "MKP-7, a Novel Mitogen-activated Protein Kinase Phosphatase,
 Functions as a Shuttle Protein";
 RL J. Biol. Chem. 276:39002-39011(2001).
 DR EMBL: AB052157; BAB47240.1; -
 DR HSSP: Q16828; IMKP.
 DR MGD: MGI:1917936; Dusp16.
 DR GO: GO:0005737; C:cytoplasm; IDA.
 DR GO: GO:0005634; C:nucleus; IDA.
 DR GO: GO:0005515; F:protein binding; IPI.
 DR GO: GO:000188; P:inactivation of MAPK; IDA.
 DR InterPro: IPR000340; DS phosphatase.
 DR InterPro: IPR000343; MAPK phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSpC; 1.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKDPHTASE.
 DR SMART: SM00195; DSpC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PSS0206; RHODANSE_3; 1.
 DR PROSITE: PSS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PSS0056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PSS0054; TYR_PHOSPHATASE_DUAL; 1.
 DR HYDROLASE: Kinase.
 SQ SEQUENCE 660 AA; 72695 MW; DB609FCDAD4AA309 CRC64;

Query Match 10.1%; Score 67; DB 11; Length 660;
 Best Local Similarity 100.0%; Pred. No. 1.5e-61;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LVALLSGREKVLILSRPVEYNTSHILEATINSGKLMKRLQDDKVLITELIOWSAR 74
 DB 15 LVALLSGREKVLILSRPVEYNTSHILEATINSGKLMKRLQDDKVLITELIOWSAR 74
 QY 75 HKVDIDC 81
 DB 75 HKVDIDC 81
 QY 81 HKVDIDC 81
 DB 75 HKVDIDC 81

RESULT 5

Q99MG6 PRELIMINARY; PRT; 677 AA.
 AC Q99MG6; (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE MAP kinase phosphatase-M A1 isoform.
 GN DUSP16 OR 3830417MI7RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RA Matsuguchi T., Musikacharoen T., Johnson T.R., Kraft A.S.,
 RA Yoshikai Y.;
 RT "A Novel MKP Is an Important Negative Regulator of LPS-mediated JNK
 Activation in Macrophages";
 RL Mol. Cell. Biol. 20:6999-7009(2001).
 DR EMBL: AF345951; AAK35052.1; -
 DR HSSP: Q16828; IMKP.
 DR MGD: MGI:1917936; Dusp16.
 DR GO: GO:0005737; C:cytoplasm; IDA.
 DR GO: GO:0005634; C:nucleus; IDA.
 DR GO: GO:0005515; F:protein binding; IPI.
 DR GO: GO:000188; P:inactivation of MAPK; IDA.

DR InterPro: IPR000340; DS phosphatase.
 DR InterPro: IPR008343; MAPK phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; Tyr_phosphatase.
 DR Pfam: PF00782; DSpC; 1.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00195; DSpC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS50206; RHODANES_3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR Hydrolase; Kinase.
 SQ SEQUENCE 677 AA; 74550 MW; 8B6DSB7096C8C2FC CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 2,3e-58;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 SFNSVHLIAGGPAEFSCFPGLCGKSTIVPTCISOPCLPVANIGPTRIIPNLVIGCOR 173
 DB 114 SFNSVHLIAGGPAEFSCFPGLCGKSTIVPTCISOPCLPVANIGPTRIIPNLVIGCOR 173

QY 174 VLNK 177
 DB 174 VLNK 177

RESULT 6
 Q99MG5 PRELIMINARY; PRT; 622 AA.

AC Q99MG5; 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Map kinase phosphatase-M A2 isoform.
 GN DUSP16 OR 3830417M17RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RA Matsuguchi T., Musikacharen T., Johnson T.R., Kraft A.S.,
 RA Yoshikai Y.;
 RT "A Novel MKP is an Important Negative Regulator of LPS-mediated JNK
 RT Activation in Macrophages";
 RL Mol. Cell. Biol. 20:6999-7009(2001).
 DR EMBL: AF345952; AAK35053.1; -.
 DR HSSP: Q16828; IMKP.
 DR MGD: MGI:1917936; Dusp16.
 DR GO: GO:0005737; Cytoplasm; IDA.
 DR GO: GO:0005634; Cytocellus; IDA.
 DR GO: GO:0005515; F:protein binding; IPI.
 DR GO: GO:0001088; P:inactivation of MAPK; IDA.
 DR InterPro: IPR000340; DS phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR000387; Tyr_phosphatase.
 DR Pfam: PF00782; DSpC; 1.
 DR Pfam: PF00581; Rhodanese; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR SMART: SM00195; DSpC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS50206; RHODANES_3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR Hydrolase; Kinase.
 SQ SEQUENCE 622 AA; 68672 MW; 6CC0A4E6B9099B98 CRC64;

Query Match
 Best Local Similarity 7.2%; Score 48; DB 11; Length 622;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCPKDPFIPESHFLRPVNDSCFCEKILPWLKSDVFIEKAKASNCVL 241
 DB 139 TCPKDPFIPESHFLRPVNDSCFCEKILPWLKSDVFIEKAKASNCVL 186

RESULT 7

Q86S58 PRELIMINARY; PRT; 625 AA.

AC Q86S58; 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to dual specificity phosphatase 8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, and Astrocytoma;
 RA Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC045110; AAH45110.1; -.
 DR GO: GO:0017017; F:MAP kinase phosphatase activity; IEA.
 DR GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro: IPR000340; DS phosphatase.
 DR InterPro: IPR008343; MAPK_phosph.
 DR InterPro: IPR002965; P-rich_extensn.
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR00387; TYR_phosphatase.
 DR Pfam: PF00782; DSpC; 1.
 DR PRINTS: PR01764; MAPKPHPTASE.
 DR PRINTS: PR01217; PRICHEXTENSN.
 DR SMART: SM00195; DSpC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS50206; RHODANES_3; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 SQ SEQUENCE 625 AA; 65826 MW; C7C808407B724FFC CRC64;

Query Match
 Best Local Similarity 3.0%; Score 20; DB 4; Length 625;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 VHCLAGISRSATIAIAYIMK 261
 DB 244 VHCLAGISRSATIAIAYIMK 263

RESULT 8

Q7TS29 PRELIMINARY; PRT; 665 AA.

AC Q7TS29; 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Dusp16 protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
 RA Brownstein M.J., Usdin T.B., Toohiyki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzyzanski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RL Strausberg R.;
 RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC052705; AAH52705.1; -
 SQ SEQUENCE 665 AA; 65021 MW; 9166E36A835249F CRC64;

Query Match 3.0%; Score 20; DB 11; Length 665;
 Best Local Similarity 100.0%; Pred. No. 1,2e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 242 VHCLAGSRGATTAATAYIMK 261
 |||||
 Db 244 VHCLAGSRGATTAATAYIMK 263

RESULT 9
 Q9YVH8 PRELIMINARY; PRT; 476 AA.
 AC Q9YVH8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE PUC protein (SD08157P).
 DE PUC OR CG7850.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V., Champs W., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Arlitt J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Berns P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Cantor A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Houtson K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mikhlov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spralling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskiy R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Maassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise B., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacleb J., Paragas V., Park S., Phuanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AEO06167; AAF54191.1; -
 DR EMBL: AY061616; AAL29164.1; -
 DR HSPF_Q16828; IMKP.
 DR FLYBase; FBgn0004210; puc.
 DR GO: GO:0006579; P:JNK kinase phosphatase activity; IMP.
 DR GO: GO:0017017; F:MAP kinase phosphatase activity; NAS.
 DR GO: GO:0030036; P:actin cytoskeleton organization and biogenesis; IMP.
 DR GO: GO:0046843; P:dorsal appendage formation; IMP.
 DR GO: GO:0008544; P:epidermal differentiation; IMP.
 DR GO: GO:0007254; P:JNK cascade; IMP.
 DR GO: GO:0046844; P:mitogen-activated protein kinase cascade; NAS.
 DR GO: GO:0046329; P:negative regulation of JNK cascade; NAS.
 DR GO: GO:0016318; P:prometastatic rotation; NAS.
 DR GO: GO:0006470; P:protein amino acid phosphorylation; IDA.
 DR GO: GO:0007396; P:protein amino acid phosphorylation; NAS.
 DR GO: GO:0042060; P:wound healing; IBP.
 DR InterPro; IPR000340; DS phosphatase.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR SMART; SM00193; DSPC; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolase.
 SQ SEQUENCE 476 AA; 51270 MW; 8E5DC281349F0F4F CRC64;

Query Match 2.0%; Score 13; DB 5; Length 476;
 Best Local Similarity 100.0%; Pred. No. 0.00024;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 246 AGISRSATTAAY 258
 |||||
 Db 218 AGISRSATTAAY 230

RESULT 10
 O46122 PRELIMINARY; PRT; 476 AA.
 ID O46122;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE PUCKERED protein.
 DE PUC OR PUCKERED OR CG7850.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

DR PROSITE; PSS0056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
 KM Hypothetical protein; Hydrolase.
 SQ SEQUENCE 657 AA; 71002 MW; A0D9153DB6326B43 CRC64;

Query Match 1.4%; Score 9; DB 5; Length 657;
 Best Local Similarity 100.0%; Pred. No. 5.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 HCLAGISRS 251
 |||||
 DB 261 HCLAGISRS 269

RESULT 13

Q82GP4 PRELIMINARY; PRT; 544 AA.
 ID Q82GP4;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative integral membrane transport protein.
 GN SAV3853.
 OS Streptomyces avermectilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxId=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shiinoe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genomic sequence of an industrial microorganism Streptomyces
 RT avermectilis: deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shiinoe M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermectilis.";
 RL Nat. Biotechnol. 21:526-531(2003).
 DR EMBL; AP005036; BAC71565.1; -.
 KM Complete proteome.
 SQ SEQUENCE 544 AA; 56291 MW; FC13A5045264A9DA CRC64;

Query Match 1.2%; Score 8; DB 16; Length 544;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 548 TSTPSTLS 555
 |||||
 DB 13 TSTPSTLS 20

RESULT 14

Q8AAB4 PRELIMINARY; PRT; 559 AA.
 ID Q8AAB4;
 AC Q8AAB4;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Asparagine synthetase B.
 GN B70551.
 OS Bacteroides thetaiotaomicron.
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 NCBI_TaxId=818;

RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=VPI-5482 / ATCC 29148;
 RX MEDLINE=22550858; PubMed=12663928;
 RA Xu J., Bjurell M.K., Himrod J., Deng S., Carmichael L.K.,
 RA Chiang H.C., Hooper L.V., Gordon J.I.;
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
 RL Science 299:2074-2076(2002).
 DR EMBL; A6016928; AAO75658.1; -.
 DR GO; GO:0004066; F:asparagine synthase (glutamine-hydrolyzing) . . .; IEA.
 DR GO; GO:0006529; P:asparagine biosynthesis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001962; Asn_synthase.
 DR InterPro; IPR006426; Asn_synth_AEB.
 DR Pfam; PF00733; Asn_synthase_1.
 DR Pfam; PF00310; GATase_2; 1.
 DR TIGRFAMs; TIGR01536; asn_synth_AEB; 1.
 KM Complete proteome.
 SQ SEQUENCE 559 AA; 63019 MW; 292F181736AD4B05 CRC64;

Query Match 1.2%; Score 8; DB 16; Length 559;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 SVPSVPSV 366
 |||||
 DB 514 SVPSVPSV 521

RESULT 15

O82K16 PRELIMINARY; PRT; 694 AA.
 ID O82K16;
 AC O82K16;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative ATP-dependent Lon protease.
 GN STM4491.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxId=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L72 / SGGC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).

DR EMBL; AEO08811; AA123309.1; -.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 KM Protease; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 694 AA; 77463 MW; 53BE7B95B5F8B878 CRC64;

Query Match 1.2%; Score 8; DB 16; Length 694;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 PVPVASEG 329
 |||||
 DB 8 PVPVASEG 15

RESULT 16

O8UR9 PRELIMINARY; PRT; 707 AA.
 ID O8UR9;
 AC O8UR9;

DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Hypothetical protein P1136.
 GN P1136.
 OS *Pyrococcus furiosus*.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the *Pyrococcus furiosus* genome."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF010223; AAL81260.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 707 AA; 80090 MW; 012ACBC9C987A6D7 CRC64;

Query Match 1.2%; Score 8; DB 17; Length 707;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 SGTETKVL 28
 |||||
 DB 53 SGTETKVL 60

RESULT 17

ID Q8YP24 PRELIMINARY; PRT; 727 AA.
 AC Q8YP24;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Hypothetical protein A114378.
 GN A114378.
 OS *Anabaena* sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=11595285; PubMed=1159840;
 RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium *Anabaena* sp. strain PCC 7120."
 RL DNA Res. 8:205-213 (2001).
 DR EMBL; AF003596; BAB76077.1; -.
 DR PIR; AB2353; AB2353.
 DR InterPro; IPR008941; TPR-like.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 727 AA; 81109 MW; 78E7D9B9FA603183 CRC64;

Query Match 1.2%; Score 8; DB 16; Length 727;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 SVPSVPSV 366
 |||||
 DB 116 SVPSVPSV 123

RESULT 18

ID Q8K058 PRELIMINARY; PRT; 858 AA.
 AC Q8K058;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

DE Hypothetical protein (Fragment).

OS *Mus musculus* (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC034090; AAH34090.1; -.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR InterPro; IPR01478; PDZ.
 DR Pfam; PF00595; PDZ; 1.
 DR PROSITE; PS50106; PDZ; 1.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 858 AA; 92849 MW; AAE908E3FA574ED CRC64;

Query Match 1.2%; Score 8; DB 11; Length 858;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 544 LAPQSTP 551
 |||||
 DB 670 LAPQSTP 677

RESULT 19

ID Q8OT06 PRELIMINARY; PRT; 892 AA.
 AC Q8OT06;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE MKIA0839 protein (Fragment).
 GN MKIA0839.
 OS *Mus musculus* (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries."
 RL DNA Res. 10:35-48 (2003).
 DR EMBL; AK122385; BAC6567.2; -.
 DR NON_TER 1
 FT NON_TER 1
 FT NON_TER 892
 SQ SEQUENCE 892 AA; 99707 MW; F2204825B83130A9 CRC64;

Query Match 1.2%; Score 8; DB 11; Length 892;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 309 SKLKLHL 316
 |||||
 DB 205 SKLKLHL 212

RESULT 20

ID Q8BMG7 PRELIMINARY; PRT; 1001 AA.
 AC Q8BMG7;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Similar to RAB3-GAP regulatory domain.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Forelimb;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT Analysis of the mouse transcriptome based on functional annotation of
 RL 60,770 full-length cDNAs.";
 DR Nature 420:563-573(2002).
 SO EMBL: AK031191; BAC27296.1; -;
 SQ SEQUENCE 1001 AA; 112208 MW; BCB82508E8201A63 CRC64;

Query Match 1.2%; Score 8; DB 11; Length 1001;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 309 SKIKLHL 316
 |||||
 DB 271 SKIKLHL 278

RESULT 21

ID Q820R6 PRELIMINARY; PRT; 1042 AA.
 AC Q820R6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Acriflavin resistance protein:heavy metal efflux pump CzCA.
 GN NE0345.
 OS Nitrosomonas europaea.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
 OC Nitrosomonadaceae; Nitrosomonas.
 NC NCB1_TaxID=915;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19718 / IFO 14298;
 RX MEDLINE=22586410; PubMed=12700255;
 RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
 RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
 RA Acleto D.M., Hommes N.G., Whitaker M.W., Arp D.J.;
 RT "Complete genome sequence of the ammonia-oxidizing bacterium and
 RL obligate chemolithoautotroph Nitrosomonas europaea.";
 DR J. Bacteriol. 185:2759-2773(2003).
 SO EMBL: BX321857; CAD84256.1; -;
 DR GO: GO:0016021; C: integral to membrane; IEA.
 DR GO: GO:0008324; F: cation transporter activity; IEA.
 DR GO: GO:0006812; P: cation transport; IEA.
 DR InterPro: IPR001036; Acrflavin_rea.
 DR InterPro: IPR004763; CzCA.
 DR Pfam: PF00873; ACR_ttrn; 1.
 DR PRINTS: PR00702; ACRIFLAVINRP.
 DR TIGRFAMs: TIGR00914; 2A0601; 1.
 KM Complete proteome.
 SQ SEQUENCE 1042 AA; 112982 MW; 471082BEC2F6B5E6 CRC64;

Query Match 1.2%; Score 8; DB 16; Length 1042;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 ALSGLHS 386
 |||||
 DB 923 ALSGLHS 930

RESULT 22
 Q815U3 PRELIMINARY; PRT; 1065 AA.
 ID Q815U3

AC Q815U3;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN PF10530C.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NC NCB1_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RX MEDLINE=22255705; PubMed=12368864;
 RA Cardon M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA Chan M.-S., Nene V., Shallow S.J., Sun B., Peterson J., Angiuoli S.,
 RA Bertea M., Allen J., Selengut J., Haft D., Mather W.W., Valdivia A.B.,
 RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carnucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrall B.,
 RT "Genome sequence of the human malaria parasite Plasmodium
 RL falciparum.";
 RL Nature 419:498-511(2002).
 DR EMBL: AE014845; AAN36195.1; -;
 KM Hypothetical protein.
 SQ SEQUENCE 1065 AA; 122856 MW; 0270CD53F55CF40 CRC64;

Query Match 1.2%; Score 8; DB 5; Length 1065;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 549 STPSLTSS 556
 |||||
 DB 325 STPSLTSS 332

RESULT 23

ID Q9FKF1 PRELIMINARY; PRT; 1194 AA.
 AC Q9FKF1;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE DNA-directed RNA polymerase subunit (EC 2.7.7.6) (DNA-directed RNA
 polymerase beta chain).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Arabidopsi; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 NC NCB1_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=98403884; PubMed=9734815;
 RA Kocani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
 RT Sequence features of the regions of 1,367,185 bp covered by 19
 RL DNA Res. 5:203-216(1998).
 CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE +
 CC {N} (N).
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
 DR EMBL: AB012240; BAB11387.1; -;
 DR GO: GO:0003677; F: DNA binding; IEA.
 DR GO: GO:0003899; F: DNA-directed RNA polymerase activity; IEA.
 DR GO: GO:0003900; F: DNA-directed RNA polymerase I activity; IEA.
 DR GO: GO:0003901; F: DNA-directed RNA polymerase II activity; IEA.
 DR GO: GO:0003902; F: DNA-directed RNA polymerase III activity; IEA.
 DR GO: GO:0016740; F: transferase activity; IEA.
 DR GO: GO:0006350; P: transcription; IEA.

```
DR InterPro: IPR007121; RNA_pol_B.
DR InterPro: IPR007644; RNA_pol_Rpb2_1.
DR InterPro: IPR007642; RNA_pol_Rpb2_2.
DR InterPro: IPR007645; RNA_pol_Rpb2_3.
DR InterPro: IPR007646; RNA_pol_Rpb2_4.
DR InterPro: IPR007647; RNA_pol_Rpb2_5.
DR InterPro: IPR007120; RNA_pol_Rpb2_6.
DR InterPro: IPR007641; RNA_pol_Rpb2_7.
DR Pfam: PF04563; RNA_pol_Rpb2_1.
DR Pfam: PF04561; RNA_pol_Rpb2_2.
DR Pfam: PF04565; RNA_pol_Rpb2_3.
DR Pfam: PF04566; RNA_pol_Rpb2_4.
DR Pfam: PF04567; RNA_pol_Rpb2_5.
DR Pfam: PF00562; RNA_pol_Rpb2_6.
DR Pfam: PF04560; RNA_pol_Rpb2_7.
DR PROSITE: PS01166; RNA_POL_BETA.1.
DR DNA-directed RNA polymerase; Transcription; Transferase.
SQ SEQUENCE 1194 AA; 134102 MW; CD48F69D16C50F8E CRC64;

Query Match 1.2%; Score 8; DB 10; Length 1194;
Best Local Similarity 100.0%; Pred.No.1.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 GTEKVLII 29
Db 203 GTEKVLII 210

RESULT 24
O01924 PRELIMINARY; PRT; 1308 AA.
ID 001924;
AC 001924;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUN-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE R155.2 protein.
GN R155.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleodermata; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RT Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RA Geisel C., Wamsley P., Kramer J.;
RT "The sequence of C. elegans cosmid R155.";
RT submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003390; AAB54272.1; -.
DR PIR: T15280; T15280.
DR WormPep: R155.2; CE12864.
DR GO: GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro: IPR000387; Tyr_phosphatase.
DR InterPro: IPR000242; Tyr_PP.
DR InterPro: IPR003125; MSN.
DR Pfam: PF02206; MSN; 1.
DR Pfam: PF00102; Y_phosphatase; 1.
DR SMART; SM00194; PTPC; 1.
DR SMART; SM00453; MSN; 1.
DR PROSITE; PSS0056; Tyr_PHOSPHATASE_2; 1.
```

```
DR PROSITE; PSS0055; TYR_PHOSPHATASE_PTP; 1.
SQ SEQUENCE 1308 AA; 146780 MW; 8FBEA85099312D CRC64;

Query Match 1.2%; Score 8; DB 5; Length 1308;
Best Local Similarity 100.0%; Pred.No.1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 247 GISRSATI 254
Db 1165 GISRSATI 1172

RESULT 25
O08TOD1 PRELIMINARY; PRT; 1885 AA.
ID O08TOD1;
AC O08TOD1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TPR-domain containing protein.
GN MAF13.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nuebaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Alnoor D., Brown A.,
RA Allen N., Naylor J., Strange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McSwan P., McKernan K., Talanas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grathame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RT Genome Res. 12:532-542(2002).
RN [2]
RP EMBL: AE010832; AAM05026.1; -.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0006396; P:RNA processing; IEA.
DR InterPro: IPR001064; Crystallin.
DR InterPro: IPR003107; HAT.
DR InterPro: IPR000504; RNA_rec_mot.
DR InterPro: IPR001440; TPR_rec_mot.
DR InterPro: IPR008941; TPR-like.
DR Pfam: PF00515; TPR; 50.
DR SMART; SM00386; HAT; 12.
DR SMART; SM00028; TPR; 48.
DR PROSITE; PSS00225; CRYSTALLIN_BETAGAMMA; 1.
DR PROSITE; PSS00030; RRM_RNP_1; 1.
DR Complete proteome.
SQ SEQUENCE 1885 AA; 214339 MW; 6D15283773777AA2 CRC64;

Query Match 1.2%; Score 8; DB 17; Length 1885;
Best Local Similarity 100.0%; Pred.No.1.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 VLGLGLK 113
Db 1840 VLGLGLK 1847

RESULT 26
O51322 PRELIMINARY; PRT; 475 AA.
ID O51322
AC O51322;
```


DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Fibronectin/fibronogen-binding protein, putative.
 OS BB0347.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
 NCBI_TaxID=139;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RA MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Karpavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Uetebach T., Matthey L., McDonald L., Artiach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RA "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi";
 RT Nature 390:580-586(1997).
 RL EMBL; AB001141; AAC6734.1; -.
 DR PIR; B70143; B70143.
 DR TIGR; BB0347; -.
 DR InterPro; IPR008653; DUF814.
 DR InterPro; IPR008616; Fbpa.
 DR Pfam; PF05670; DUF814; 1.
 DR Pfam; PF05833; Fbpa; 1.
 KW Complete proteome.
 SQ SEQUENCE 475 AA; 55912 MW; 94E07528DB6EB4 CRC64;

Query Match 1.1%; Score 7; DB 16; Length 475;
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 KXIKNOT 302
 DB 469 KXIKNOT 475

RESULT 27
 Q9W131 PRELIMINARY; PRT; 478 AA.
 AC 09W131;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE CG13594 protein.
 GN CG13594.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballow R.M., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasbo P., Lei Y., Levitky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Modyarty C., Morris J., Mostreft A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munzy D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacble J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Sanders R.D.C., Scheller F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195 (2000).
 [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon K.Y., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorese V., Doup L.E., Doyle C., Drenek D., Fartan D.,
 RA Ferreira S., Friese E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Mostreft A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacble J., Paragae V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheller F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RA Mistra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Friese E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF003464; AAF4246.2; -.
 DR FlyBase; FBgn035041; CG13594.
 SQ SEQUENCE 478 AA; 49256 MW; 3C56261A6EB8F8A CRC64;

Query Match 1.1%; Score 7; DB 5; Length 478;
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 YVLANSY 193
 DB 19 YVLANSY 25

```

RESULT 28
Q9S3Z3      PRELIMINARY;      PRT;      479 AA.
AC  Q9S3Z3;
DT  01-MAY-2000 (T-EMBLrel. 13, Created)
DT  01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DE  01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE  Acyl-CoA reductase LuxC.
GN  LUXC.
OS  Vibrio fischeri.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC  Vibrionaceae; Vibrio.
OX  NCBI_TaxID=668;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=MJ-1;
RA  "Knight T., Papadakis N.;
RT  "Vibrio fischeri lux operon SalI digest.";
RL  Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF170104; AAD48475.1; -
DR  GO; GO:0019109; F:acyl-CoA reductase activity; IEA.
DR  GO; GO:0008218; P:bioluminescence; IEA.
DR  InterPro; IPR008670; LuxC.
DR  Pfam; PF05893; LuxC; 1.
SQ  SEQUENCE 479 AA; 54578 MW; 10FD9567E5026340 CRC64;

Query Match      1.1%; Score 7; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  117 SVHLIAG 123
DB  146 SVHLIAG 152

RESULT 29
Q9LZG8      PRELIMINARY;      PRT;      479 AA.
AC  Q9LZG8;
DT  01-OCT-2000 (T-EMBLrel. 15, Created)
DT  01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT  01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE  Transporter-like protein.
GN  T8A8_80.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Parnelle B., Boutry M., Goffeau A., Mewes H.W., Lemcke K.,
RA  Mayer K.F.X., Quelet F., Salanoubat M.;
RL  Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RA  EU Arabidopsis sequencing project;
RL  Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC  -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR  EMBL; AL162691; CAB83151.1; -.
DR  PIR; T47415; T47415.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0005215; F:transporter activity; IEA.
DR  GO; GO:0006810; P:transport; IEA.
DR  InterPro; IPR007114; MFS.
DR  InterPro; IPR005828; Sub transporter.
DR  Pfam; PF00063; sugar tr; 1.
DR  PROSITE; PSS0850; MFS; 1.
DR  Transmembrane.
SQ  SEQUENCE 479 AA; 52806 MW; 8344F896FB83ECA CRC64;

Query Match      1.1%; Score 7; DB 10; Length 479;

```

```

Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  518 LFGISTS 524
DB  117 LFGISTS 123

RESULT 30
Q56817      PRELIMINARY;      PRT;      480 AA.
AC  Q56817;
DT  01-NOV-1996 (T-EMBLrel. 01, Created)
DT  01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT  01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE  Fatty acid reductase.
GN  LUXC.
OS  Photorhabdus luminescens (Xenorhabdus luminescens).
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC  Enterobacteriaceae; Photorhabdus.
OX  NCBI_TaxID=29488;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=HB;
RX  MEDLINE=9235513; PubMed=1644764;
RA  "Meighen E.A., Sziltnner R.;
RT  "Multiple repetitive elements organization of the lux operons of
RT  luminescent terrestrial bacteria.";
RL  J. Bacteriol. 174:5371-5381(1992).
DR  EMBL; M90093; AAA27617.1; -
DR  GO; GO:0019109; F:acyl-CoA reductase activity; IEA.
DR  GO; GO:0008218; P:bioluminescence; IEA.
DR  InterPro; IPR008670; LuxC.
DR  Pfam; PF05893; LuxC; 1.
SQ  SEQUENCE 480 AA; 54636 MW; AD42A2BF091B2B3 CRC64;

Query Match      1.1%; Score 7; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  117 SVHLIAG 123
DB  147 SVHLIAG 153

RESULT 31
Q8SY11      PRELIMINARY;      PRT;      480 AA.
AC  Q8SY11;
DT  01-JUN-2002 (T-EMBLrel. 21, Created)
DT  01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT  01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE  REL6208p.
GN  CG13594.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxID=7227;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Berkely;
RA  Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA  Champe M., Chavez C., Dorsett V., Drensek D., Farfan D., Frise B.,
RA  George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA  Miranda A., Mungall C.J., Nunoo J., Pacleb J., Pargass V., Park S.,
RA  Patel S., Phoumenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA  Celinker S.;
RL  Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AY075460; AAL68273.1; -.
DR  FlyBase; FBgn0035041; CG13594.
SQ  SEQUENCE 480 AA; 49470 MW; 052E1B3317EDD700 CRC64;

```

Query Match 1.1%; Score 7; DB 5; Length 480;
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 YVUNASY 193
 |||||
 DB 21 YVUNASY 27

RESULT 32

Q9H6A7 PRELIMINARY; PRT; 483 AA.
 AC Q9H6A7;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein FLJ22429.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 ON NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawabata A., Hiki J. T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isegaki T., Sugano S.;
 RT "NEDO human cDNA sequencing project."
 RU Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA Strauberg R.;
 RU Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK026882; BAB15355.1; -.
 DR EMBL; BC012837; AAH12837.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 483 AA; 53266 MW; C16B129B6005BAAF CRC64;

Query Match 1.1%; Score 7; DB 4; Length 483;
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 487 VRTSSSG 493
 |||||
 DB 53 VRTSSSG 59

RESULT 33

Q9SXY1 PRELIMINARY; PRT; 487 AA.
 ID Q9SXY1;
 AC Q9SXY1;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE FAS2 (Fragment).
 GN FAS2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 ON NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RA Arai T., Kaya H.;
 RT "FASCIAT2, open reading frame."
 RU Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=98290546; PubMed=9628582;
 RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
 RA Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
 RT Sequence features of the regions of 1,456,315 bp covered by nineteen
 RT physically assigned p1 and TAC clones.";

RL DNA Res. 5:41-54(1998).
 DR EMBL; AB027230; BAB77766.1; -.
 DR EMBL; AB010076; BAB11430.1; -.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 5.
 DR SMART; SM00320; WD40; 5.
 DR PROSITE; PS00678; WD_REPEATS_1; 1.
 DR PROSITE; PS00682; WD_REPEATS_2; 3.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT NON TER 487
 SQ SEQUENCE 487 AA; 54106 MW; 06CB34C29E1F3191 CRC64;

Query Match 1.1%; Score 7; DB 10; Length 487;
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 VASLSSD 101
 |||||
 DB 169 VASLSSD 175

RESULT 34

Q8ZJT2 PRELIMINARY; PRT; 488 AA.
 ID Q8ZJT2;
 AC Q8ZJT2;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein YPO0004.
 GN YPO0004 OR Y0004.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 ON NCBI_Taxid=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Tiddall R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltham D., Hamlin N., Holtroyd S., Jags K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyelson P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague."
 RU Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIM5 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Lies P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of Yersinia pestis KIM.";
 RU J. Bacteriol. 184:4601-4611(2002).
 DR EMBL; A014141; CAC88871.1; -.
 DR EMBL; A013601; AAM83600.1; -.
 DR PIR; AE0001; AE0001.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 488 AA; 56081 MW; CD17471A7A85D82D CRC64;

Query Match 1.1%; Score 7; DB 16; Length 488;
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SLEEDSP 375

DB 80 SLEDS 86

RESULT 35

O8IMG2

ID O8IMG2 PRELIMINARY; PRT; 490 AA.

AC O8IMG2; 01-JUN-2003 (T-EMBLrel. 24, Created)

DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE Stage V sporulation protein AF.

GN SPOVAF OR BA4286.

OS Bacillus anthracis (strain Ames).

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=198094;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22608414; PubMed=12721629;

RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,

RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,

RA Holtapple E.K., Okstad O.A., Helgason E., Ristic J., Wu M.,

RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,

RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,

RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,

RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,

RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,

RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,

RA Thomas B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,

RA Fraser C.M.;

RT "The genome sequence of *Bacillus anthracis* Ames and comparison to

RT closely related bacteria.";

RL Nature 423:81-86(2003).

DR EMBL; AB017037; AAP28006.1; -.

DR TIGR; BA4286; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0009847; P:spore germination; IEA.

DR InterPro; IPR004995; GerA.

DR Pfam; PF03323; GerA; 1.

KW Complete proteome.

SQ SEQUENCE 490 AA; 55064 MW; 02EBF57EB86009C CRC64;

Query Match 1.1%; Score 7; DB 16; Length 490;

Best Local Similarity 100.0%; Pred. No. 5.9e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 377 VQALSGI 383

DB 104 VQALSGI 110

RESULT 36

O9KXD8 PRELIMINARY; PRT; 492 AA.

AC O9KXD8; 01-OCT-2000 (T-EMBLrel. 15, Created)

DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE Hypothetical protein.

GN H0115.

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=83334;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20198780; PubMed=10734605;

RA Makino K., Yokoyama K., Kubota Y., Yutsudo C.H., Kimura S.,

RA Kurokawa K., Ishii K., Hattori M., Tatsuno I., Abe H., Iida T.,

RA Yamamoto K., Ohnishi M., Hayashi T., Yasunaga T., Honda T.,

RA Sasakawa C., Shinagawa H.;

RT "Complete nucleotide sequence of the prophage VT2-Sakai carrying the

RT verotoxin 2 genes of the enterohemorrhagic *Escherichia coli* O157:H7

RT derived from the Sakai outbreak.";

RL Genes Genet. Syst. 74:227-239(1999).

DR EMBL; AF000422; BA94143.1; -.

DR InterPro; IPR005181; DUF303.

DR Pfam; PF03629; DUF303; 1.

KW Hypothetical protein.

SQ SEQUENCE 492 AA; 52888 MW; 42BEFA7F2274B0DC CRC64;

Query Match 1.1%; Score 7; DB 2; Length 492;

Best Local Similarity 100.0%; Pred. No. 5.9e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 417 LHGFSSS 423

DB 475 LHGFSSS 481

RESULT 37

O8ZYH4 PRELIMINARY; PRT; 494 AA.

AC O8ZYH4; 01-MAR-2002 (T-EMBLrel. 20, Created)

DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE Conserved protein (Possible type II secretion, gsp).

GN PA0773.

OS *Pyrobaculum aerophilum*.

OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;

OC Thermoproteaceae; Pyrobaculum.

OX NCBI_TaxID=13773;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=IM2 / ATCC 51768 / DSM 7523;

RX MEDLINE=2164397; PubMed=11792869;

RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,

RA Miller J.H.;

RT "Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum*

RT aerophilum.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).

DR EMBL; AE009784; AAL63019.1; -.

DR GO; GO:0005622; C:intracellular; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0008170; F:N-methyltransferase activity; IEA.

DR GO; GO:0006306; F:DNA methylation; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR001482; GSP1_E.

DR InterPro; IPR002052; N6_Mtase.

DR Pfam; PF00437; GSP1_E; 1.

DR Pfam; PF000739; GSP1_E; 1.

DR PROSITE; PS00092; N6_MTASE; 1.

KW Complete proteome.

SQ SEQUENCE 494 AA; 56443 MW; D70B2A55341B0B41 CRC64;

Query Match 1.1%; Score 7; DB 17; Length 494;

Best Local Similarity 100.0%; Pred. No. 5.9e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 SREELGK 648

DB 69 SREELGK 75

RESULT 38

O9W0U3 PRELIMINARY; PRT; 495 AA.

AC O9W0U3; 01-MAY-2000 (T-EMBLrel. 13, Created)

DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)

DE CG7042 protein.

GN CG7042.

OS *Drosophila melanogaster* (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wen K.H., Doyle C.C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abri J.F., Agbayani C., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Paoli B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Gloddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodali C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spter E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirbas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RA Celinker S.E., Adams M.D., Krommler B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fafan D.,
 RA Ferriere S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ileguam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuno J.,
 RA Pacle J., Parag V., Park S., Patel S., Pfeiffer B.,
 RA Phouanouvong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirbas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celinker S.E.,
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommler B., Marshall B., Millburn G., Richer J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutnack F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;

RT "Annotation of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; A6003467; AAF47341.2; -;
 DR FlyBase; FBgn0035105; CG7042.
 DR GO; GO:0008138; F:protein tyrosine/serine/threonine phosphatase. . . IEA.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR001406; Ds phosphatase.
 DR InterPro; IPR001406; Pseudosynth_1.
 DR InterPro; IPR000387; Tyr_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR Pfam; PF01416; Pseudosynth_1; 1.
 DR SMART; SM00195; DSPC; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS00386; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 SQ SEQUENCE 495 AA; 56479 MW; 9A0BDE2A42069C3 CRC64;
 Query Match 1.1%; Score 7; DB 5; Length 495;
 Best Local Similarity 100.0%; Pred. No. 5; 9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 238 GCYLVHC 244
 Db 438 GCYLVHC 444
 RESULT 39
 ID 076107 PRELIMINARY; PRT; 499 AA.
 AC 076107;
 DT 01-NOV-1998 (TEMBLrel. 08, Created)
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Cytochrome P450.
 GN CYP6B1.
 OS *Culex quinquefasciatus* (Southern house mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culex.
 OC NCBI_TaxID=7176;
 OX NCBI_TaxID=7176;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JPAL-per; TISSUE=gut;
 RX MEDLINE=98194786; PubMed=9535163;
 RA Kasai S., Shono T., Yamakawa M.;
 RT "Molecular cloning and nucleotide sequence of a cytochrome P450 cDNA
 from a pyrethroid-resistant mosquito, *Culex quinquefasciatus* Say.";
 RL Insect Mol. Biol. 7:185-190(1998).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; A6001323; BAA28946.1; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
 DR GO; GO:0006306; P:DNA methylation; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR002052; N6_Mtase.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 DR PROSITE; PS00092; N6_MTASE; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 499 AA; 57486 MW; A1C42A04E3ECF238 CRC64;
 Query Match 1.1%; Score 7; DB 5; Length 499;
 Best Local Similarity 100.0%; Pred. No. 6e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 544 LAPOTST 550
|||||
Db 455 LAPOTST 461

RESULT 40

Q9N3R4 PRELIMINARY; PRT; 502 AA.
AC Q9N3R4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN Y47G6A.1.
OS Caenorhabditis elegans.
OC Bkaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Minx P., Graves T., Hawrysko C.;
RT "The sequence of C. elegans cosmid Y47G6A.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024791; AAF60654.2; -.
DR WormBep; Y47G6A.1; CE27258.
DR InterPro; IPR000990; Innexin.
DR Pfam; PF00876; Innexin; 1.
DR PRINTS; PR01262; INNEXIN.
KW Hypothetical protein.
SQ SEQUENCE 502 AA; 57540 MW; 29178A4D71580D0D CRC64;

Query Match 1.1%; Score 7; DB 5; Length 502;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LBSGTEK 25
|||||
Db 40 LBSGTEK 46

RESULT 41

O806Y8 PRELIMINARY; PRT; 505 AA.
AC O806Y8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA-dependent RNA polymerase (Fragment).
GN RDRP.
OS Cucumber yellows virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
OC Citrinivirus.
OX NCBI_TaxID=32618;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541874; PubMed=12655104;

RA Hartono S., Natsunaki T., Genda Y., Okuda S.;
RT "Nucleotide sequence and genome organization of Cucumber yellows virus, a member of the genus Citrinivirus.";
RL J. Gen. Virol. 84:1007-1012(2003).
DR EMBL; AB085613; BAC66371.1; -.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.

DR InterPro; IPR001788; RNA_dep_RNAPol2.
DR InterPro; IPR007095; RNA_pol_DS_Ps.
DR InterPro; IPR007094; RNA_pol_PsVlr.
DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
KW RNA-directed RNA polymerase.
FT NON TER 1
SQ SEQUENCE 505 AA; 58437 MW; 877D9E228B68DB5 CRC64;

Query Match 1.1%; Score 7; DB 12; Length 505;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 395 KTKRSFS 401
|||||
Db 166 KTKRSFS 172

RESULT 42

Q7UH12 PRELIMINARY; PRT; 508 AA.
AC Q7UH12;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P0P1-Artpase beta chain.
GN ATPD OR R84906.
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Glockner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heltmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp. strain 1.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294141; CAD78167.1; -.
KW Complete proteome.

SQ SEQUENCE 508 AA; 55235 MW; F4E4A544A48BCFB9 CRC64;
Query Match 1.1%; Score 7; DB 16; Length 508;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 VLITELI 69
|||||
Db 210 VLITELI 216

RESULT 43

O50188 PRELIMINARY; PRT; 509 AA.
AC O50188; O50187;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Probable phosphoprotein phosphatase.
GN PPP.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 [1]
 RP SEQUENCE FROM N.A.
 RA Cole S.T.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9318700; PubMed=8446027;
 RA Eiglmeier K., Honore N., Woods S.A., Caudron B., Cole S.T.;
 RT "Use of an ordered cosmid library to deduce the genomic organization
 of Mycobacterium leprae."
 RL Mol. Microbiol. 7:197-206(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97124199; PubMed=8969512;
 RA Fasih H., De Rossi E., Salazar L., Cantoni R., Labo M., Ricciardi G.,
 Takiff H.E., Eiglmeier K., Bergh S., Cole S.T.;
 RT "Gene arrangement and organization in a approximately 76 kb fragment
 encompassing the orfC region of the chromosome of Mycobacterium
 leprae."
 RL Microbiol. 142:3147-3161(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96310367; PubMed=8733228;
 RA Salazar L., Fasih H., De Rossi E., Ricciardi G., Rios C., Cole S.T.,
 Takiff H.E.;
 RT "Organization of the origins of replication of the chromosomes of
 Mycobacterium smegmatis, Mycobacterium leprae and Mycobacterium
 tuberculosis and isolation of a functional origin from M. smegmatis."
 RL Mol. Microbiol. 20:283-293(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96194983; PubMed=8622949;
 RA Fasih H., Vincent V., Cole S.T.;
 RT "Homing events in the gyrA gene of some mycobacteria."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:3410-3415(1996).
 DR EMBL: Z70722; CAA94721.1; -
 DR PIR: T10013; T10013
 DR GO: 0003824; F:catalytic activity; IEA.
 DR InterPro: IPR001932; PP2C-like.
 DR InterPro: IPR003643; Pto_Pase.
 DR Pfam: PF00481; PP2C; 1.
 DR Prodom: PD006823; Pto_Pase; 1.
 DR SMART: SM00332; PP2C_1.
 DR SMART: SM00331; PP2C_SIG; 1.
 DR SEQUENCE 509 AA; 53661 MW; 7D8DC9691186DF9C CRC64;
 SQ
 Query Match 1.1%; Score 7; DB 2; Length 509;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 489 TSSSGTA 495
 |||||
 Db 460 TSSSGTA 466
 RESULT 44
 Q9CDE5
 AC Q9CDE5; PRELIMINARY; PRT; 509 AA.
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein ML0020.
 GN ML0020.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TW;

RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
 Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R.M., Devlin K., Duthoy S., Fellwell T., Fraser A., Hamlin N.,
 Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 Barrett B.G.;
 RT "Massive gene decay in the leprosy bacillus."
 RL Nature 409:1007-1011(2001).
 DR EMBL: AL583917; CAC29528.1; -
 DR PIR: D86911; D86911.
 DR Leproma: ML0020; -
 DR GO: 0003824; F:catalytic activity; IEA.
 DR InterPro: IPR001932; PP2C-like.
 DR InterPro: IPR003643; Pto_Pase.
 DR Pfam: PF00481; PP2C; 1.
 DR Prodom: PD006823; Pto_Pase; 1.
 DR SMART: SM00332; PP2C_1.
 DR SMART: SM00331; PP2C_SIG; 1.
 DR SMART: SM00331; PP2C_SIG; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 509 AA; 53576 MW; 6D0B3DF76D9D1ED9 CRC64;
 Query Match 1.1%; Score 7; DB 16; Length 509;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 489 TSSSGTA 495
 |||||
 Db 460 TSSSGTA 466
 RESULT 45
 Q8YGD2
 ID Q8YGD2; PRELIMINARY; PRT; 511 AA.
 AC Q8YGD2;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein BME11227.
 GN BME11227.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mujer C., Los T.,
 Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 Jablonksi L., Larsen N., D'Souza M., Bernal A., Mazur M., Gotsman E.,
 Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 Haselkorn R., Kyrides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009562; AL52408.1; -
 DR PIR: AE3405; AE3405.
 DR GO: 0008080; F:N-acetyltransferase activity; IEA.
 DR InterPro: IPR00182; GCN5acetyl_trans.
 DR Pfam: PF00583; Acetyltransf; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 511 AA; 57832 MW; 729C9855EBB414F8 CRC64;
 Query Match 1.1%; Score 7; DB 16; Length 511;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 73 AKHKVDI 79
 |||||

Db 34 AKHKVDI 40

RESULT 46

ID 08GIJ0 PRELIMINARY; PRT; 511 AA.

AC 08GIJ0; (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

GN Acetyltransferase, GNAT family.

OS BR0725.

OC Brucella suis.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Brucellaceae; Brucella.

OX NCB1_TaxID=29461;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1330 / Biovar 1;

RX MEDLINE=22247741; PubMed=12271122;

RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,

RA Read T.D., Dodson R.V., Umayam L., Brinkac L.M., Beanan M.J.,

RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,

RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,

RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,

RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;

RT "The Brucella suis genome reveals fundamental similarities between

animal and plant pathogens and symbionts."

RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).

DR EMBL; AE014378; AAN29654.1; -.

DR TIGR; BR0725; -.

DR GO; GO:0008080; P:N-acetyltransferase activity; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR InterPro; IPR000182; GCN5acetyl_trans.

DR Pfam; PF00583; Acetyltransf. 1.

DR TRANSFAC; Complete proteome.

SQ SEQUENCE 511 AA; 57628 MW; 411004CA35883AAE CRC64;

Query Match 1.1%; Score 7; DB 16; Length 511;

Best Local Similarity 100.0%; Pred. No. 6.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 AKHKVDI 79

ID 08GIJ0 PRELIMINARY; PRT; 515 AA.

AC 08GIJ0; (TREMBlrel. 02, Created)

DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)

DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)

DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE 3-ketosteroid-1-dehydrogenase (EC 1.3.99.4).

GN KSDP.

OS Nocardioides simplex (Arthrobacter simplex).

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Propionibacteriaceae; Nocardioidaceae; Pimelobacter.

OX NCB1_TaxID=2045;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=IFO12096;

RX MEDLINE=95319331; PubMed=7596291;

RA Molnar I., Choi K., Yamashita M., Murooka Y.;

RT "Molecular cloning, expression in Streptomyces lividans, and analysis

of a gene cluster from Arthrobacter simplex encoding 3-

ketosteroid-DELTA.1-dehydrogenase, 3-ketosteroid-DELTA.5-isomerase

RT and a hypothetical regulatory protein."

RL Mol. Microbiol. 15:895-905(1995).

DR EMBL; D37969; BAA07186.1; -.

DR PIR; S61889; S61889.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR005829; Sug_transporter.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.

KW Oxidoreductase.

SQ SEQUENCE 515 AA; 54330 MW; 78CFECA65AB022DB CRC64;

Query Match 1.1%; Score 7; DB 2; Length 515;

Best Local Similarity 100.0%; Pred. No. 6.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 EDALERY 430

Db 71 EDALERY 77

RESULT 48

ID 082NP2 PRELIMINARY; PRT; 516 AA.

AC 082NP2; (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Putative acyl-CoA synthetase, long-chain fatty acid:CoA ligase.

GN FAD06 OR SAV1259.

OS Streptomyces avermitilis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomyces; Streptomyces; Streptomycetaceae; Streptomyces.

OX NCB1_TaxID=33903;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RX MEDLINE=21477403; PubMed=11572948;

RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,

RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;

RT "Genome sequence of an industrial microorganism Streptomyces

RT avermitilis: deducing the ability of producing secondary

metabolites."

RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RX MEDLINE=22608306; PubMed=12692562;

RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,

RA Sakaki Y., Hattori M., Omura S.;

RT "Complete genome sequence and comparative analysis of the industrial

microorganism Streptomyces avermitilis."

RL Nat. Biotechnol. 21:526-531(2003).

DR EMBL; AP005026; BAC68969.1; -.

DR GO; GO:0016874; F:ligase activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR000873; AMP-bind.

DR Pfam; PF00501; AMP-binding; 1.

DR PROSITE; PS00455; AMP BINDING; 1.

KW ligase; Complete proteome.

SQ SEQUENCE 516 AA; 56513 MW; 6255B86728639A6A CRC64;

Query Match 1.1%; Score 7; DB 16; Length 516;

Best Local Similarity 100.0%; Pred. No. 6.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 KSVDFIE 231

Db 483 KSVDFIE 489

RESULT 49

ID 093N44 PRELIMINARY; PRT; 519 AA.

AC 093N44; (TREMBlrel. 19, Created)

Search completed: June 21, 2004, 13:28:28
Job time : 52 secs

DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Sulfotransferase (Hypothetical protein).
GN CB06699.
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; Coxiella.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RA Hoover T.A., Vodkin M.H., Williams J.C., Culp D.W., Thompson H.A.;
RT "A chromosomal DNA deletion explains the phenotype of the Coxiella
burnetii phase II variant."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile phase I / RSA 493;
RX MEDLINE=22608657; PubMed=12704232;
RA Seeshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA Nelson W.C., Ward N.L., Tettelin H., Davidson T.M., Beanan M.J.,
RA DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
RT "Complete genome sequence of the Q-fever pathogen, Coxiella
burnetii."
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
DR EMBL: AF387640; AAK71278.1; -.
DR EMBL: AE016862; AAO90243.1; -.
DR TIGR: CB00699; -.
DR GO: 0016740; F:transferase activity; IRA.
KW Hypothetical protein; Transferase; Complete proteome.
SQ SEQUENCE 519 AA; 61647 MW; 9D6E70B81AC690C4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 16; Length 519;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 LEANIN 49
| | | | |
DB 36 LEANIN 42

RESULT 50
Q8N3D9
ID Q8N3D9 PRELIMINARY; PRT; 523 AA.
AC Q8N3D9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP762F135.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RA Bloembergen H., Boecker M., Brandt P., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL834428; CAD39089.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 523 AA; 54448 MW; 0B58F308E13320F8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 4; Length 523;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 HPASVPS 362
| | | | |
DB 341 HPASVPS 347

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:21:13 ; Search time 58 Seconds
(without alignments)
3239.553 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665

Sequence: 1 MAHEMIGTIVTERLVALLE.....LKVSGSQSFGSGMEIIEVS 665

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 162285

Minimum DB seq length: 473

Maximum DB seq length: 200000000

Post-Processing: Listing first 75 summaries

Database : A_Geneseq_29yand4:*

1: geneseq1980s:***
2: geneseq1990s:***
3: geneseq2000s:***
4: geneseq2001s:***
5: geneseq2002s:***
6: geneseq2003as:***
7: geneseq2003bs:***
8: geneseq2004s:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	665	100.0	665	5 ABR52381	Abt52381 Protein r
2	665	100.0	665	5 ABR52407	Abt52407 Protein r
3	472	71.0	517	5 AAU79159	AAU79159 Human dua
4	472	71.0	665	4 AAEO4834	AAEO4834 Human SGP
5	472	71.0	665	4 AAU09016	AAU09016 Human dua
6	472	71.0	665	4 AAU79156	AAU79156 Human dua
7	472	71.0	665	5 AAU09946	AAU09946 Protein s
8	472	71.0	665	5 AAU75789	AAU75789 Human pro
9	472	71.0	665	5 ABB97946	ABB97946 Human pro
10	472	71.0	665	5 AAU79929	AAU79929 Human dua
11	472	71.0	665	5 ABB97291	ABB97291 Novel hum
12	472	71.0	665	6 ADA54744	ADA54744 Human pro
13	472	68.0	665	5 AAU79161	AAU79161 Human dua
14	421	63.3	665	5 AAU79162	AAU79162 Human dua
15	350	52.6	665	5 ABR52352	ABR52352 Protein r
16	281	42.3	664	5 ABR52424	ABR52424 Protein r
17	266	40.0	666	4 AAB20325	AAB20325 Human pro
18	258	38.8	672	4 AAM25744	AAM25744 Human pro
19	64	9.6	660	5 ABR52385	ABR52385 Protein r
20	20	3.0	579	7 ADE08458	ADE08458 Novel pro
21	20	3.0	625	5 ABR52382	ABR52382 Protein r
22	20	3.0	625	5 ABR52350	ABR52350 Protein r
23	20	3.0	625	6 ABG73440	ABG73440 Human dua
24	20	3.0	663	2 AAW29150	AAW29150 dual-spec
25	20	3.0	663	5 ABR52351	ABR52351 Protein r

26	13	2.0	476	4 ABB63527	ABB63527 Drosophila
27	12	1.8	436	4 ABG00724	ABG00724 Novel hum
28	8	1.2	876	7 ADA44807	ADA44807 CD4/TCR C
29	8	1.2	474	3 AAY59170	AAY59170 CD4-Ig fu
30	8	1.2	481	1 AAP93011	AAP93011 Genetic c
31	8	1.2	481	3 AAB19510	AAB19510 CD4-IgM f
32	8	1.2	481	3 AAY51081	AAY51081 Human fus
33	8	1.2	481	3 AAY59171	AAY59171 CD4-Ig fu
34	8	1.2	507	3 AAB00158	AAB00158 bCD4-SCFv
35	8	1.2	519	2 AAR20152	AAR20152 Human CD4
36	8	1.2	524	1 AAP94703	AAP94703 Sequence
37	8	1.2	530	2 AAR26783	AAR26783 CD4-IgG2
38	8	1.2	530	2 AAR46679	AAR46679 CD4-IgG2
39	8	1.2	530	3 AAY85080	AAY85080 CD4-IgG2
40	8	1.2	530	4 AAB67323	AAB67323 CD4-IgG2
41	8	1.2	530	4 AAB80864	AAB80864 Human CD4
42	8	1.2	530	6 ABG71123	ABG71123 CD4-Immuna
43	8	1.2	532	2 AAR27278	AAR27278 CD4-gamma
44	8	1.2	532	2 AAR78678	AAR78678 T-cell re
45	8	1.2	532	2 AAR89458	AAR89458 CD4-eta f
46	8	1.2	532	2 AAW02215	AAW02215 CD4-T-cel
47	8	1.2	532	2 AAW83141	AAW83141 Chimeric
48	8	1.2	534	2 AAR26531	AAR26531 Sequence
49	8	1.2	549	2 AAR04920	AAR04920 Immunopro
50	8	1.2	557	2 AAR04923	AAR04923 Immunopro
51	8	1.2	574	2 AAR04919	AAR04919 Immunopro
52	8	1.2	575	2 AAR27276	AAR27276 CD4-zeta
53	8	1.2	575	2 AAR78676	AAR78676 T-cell re
54	8	1.2	575	2 AAR89456	AAR89456 CD4-zeta
55	8	1.2	575	2 AAW02213	AAW02213 CD4-T-cel
56	8	1.2	575	2 AAW83140	AAW83140 Chimeric
57	8	1.2	577	2 AAR04924	AAR04924 Immunopro
58	8	1.2	590	6 ABU07697	ABU07697 Viral coa
59	8	1.2	614	1 AAP93012	AAP93012 Genetic c
60	8	1.2	616	3 AAB19511	AAB19511 CD4-IgG1
61	8	1.2	616	3 AAY51082	AAY51082 Human fus
62	8	1.2	616	3 AAY59172	AAY59172 CD4-Ig fu
63	8	1.2	631	1 AAP93009	AAP93009 Genetic c
64	8	1.2	631	3 AAB19508	AAB19508 CD4-IgG1
65	8	1.2	631	3 AAY51079	AAY51079 Human fus
66	8	1.2	631	3 AAY59169	AAY59169 CD4-Ig fu
67	8	1.2	671	3 AAW68500	AAW68500 Hybrid re
68	8	1.2	720	6 ABU07694	ABU07694 Viral coa
69	8	1.2	720	6 ABU07695	ABU07695 Viral coa
70	8	1.2	729	1 AAP93008	AAP93008 Genetic c
71	8	1.2	729	2 AAR41042	AAR41042 CD4-GBPH
72	8	1.2	729	3 AAB19507	AAB19507 CD4-IgG1
73	8	1.2	729	3 AAY51078	AAY51078 Human fus
74	8	1.2	729	3 AAY59168	AAY59168 CD4-Ig fu
75	8	1.2	903	2 AAR60183	AAR60183 PA(1-725)

ALIGNMENTS

RESULT 1
ABR52381
ID ABR52381 standard; protein; 665 AA.
XX
AC ABR52381;
XX
DT 19-UTN-2003 (first entry)
XX
DE Protein relating to the invention SEQ ID NO: 109.
XX
KW antiproliferative; hepatotropic; nephrotropic; antiarrhythmic;
KW antiproliferative; cardiac; cytosolic; gene therapy; liver disease;
KW proliferative disorder; renal failure; cardiovascular disorder;
KW immunological disorder; arthritis; psoriasis; congenital heart defect;
KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
OS Homo sapiens.
XX

PN WO200257460-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 20-DEC-2001; 2001WO-US050459.
 XX
 PF 20-DEC-2000; 2000US-0256868P.
 PR 30-MAR-2001; 2001US-0280186P.
 PR 01-MAY-2001; 2001US-0287735P.
 PR 05-JUN-2001; 2001US-0295848P.
 PR 25-JUN-2001; 2001US-0300465P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
 PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
 PI Krycek S, Mcatee P, Suchard S, Banas D;
 DR WPI; 2002-599721/64.
 DR N-PSDB; ACC60559.
 XX
 PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 PS Claim 5; Fig 12; 801pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid comprising a
 CC polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotides fully defined in the specification. The polynucleotide of
 CC the invention has antiproliferative, hepatotropic, nephrotropic,
 CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
 CC polynucleotide may have a use in gene therapy. A polynucleotide or
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention
 XX
 SQ Sequence 665 AA;
 Query Match 100.0%; Score 665; DB 5; Length 665;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQO 60
 DB 1 MAHEMIGTQIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRLQO 60
 QY 61 DKVLTTELIOHSAKRVKIDCSQKVVYVYQSSQDVASLSSDQFLVYLKLEKESFNHYL 120
 DB 61 DKVLTTELIOHSAKRVKIDCSQKVVYVYQSSQDVASLSSDQFLVYLKLEKESFNHYL 120
 QY 121 IAGFAEFRCPPGCEGKSTLVPTCISOPCLPVANIGTRTILPMUYGCGRDVANKKELI 180
 DB 121 IAGFAEFRCPPGCEGKSTLVPTCISOPCLPVANIGTRTILPMUYGCGRDVANKKELI 180
 QY 181 QONGIGYVLNASTYCPKDPFIPESHFLRPVNDSECEKILPWLIDKSVDPIERAKASNGCV 240
 DB 181 QONGIGYVLNASTYCPKDPFIPESHFLRPVNDSECEKILPWLIDKSVDPIERAKASNGCV 240
 QY 241 LVHCLAGISRSKATIAIYMKMDMSLDBAYFVVEKRPITSPNNFGLQILDIYEKIKN 300
 DB 241 LVHCLAGISRSKATIAIYMKMDMSLDBAYFVVEKRPITSPNNFGLQILDIYEKIKN 300
 QY 301 QTGAGGPPSKKLTLLHEKNEPVPVASEGQKSEFTPIASPPCADTSSEAAGRPVHPASV 360
 DB 301 QTGAGGPPSKKLTLLHEKNEPVPVASEGQKSEFTPIASPPCADTSSEAAGRPVHPASV 360
 QY 361 PSVPVQPSLLEDSPLVQALSGHLASDRLEDSNKLKRSFSLDIKSVSYASMAASLHG 420

DB 361 PSVPVQPSLLEDSPLVQALSGHLASDRLEDSNKLKRSFSLDIKSVSYASMAASLHG 420
 QY 421 SSSDALRYKRPSTLLDGTNKLCOFSPVOELSEQPTSPDKEKASIRKLOTAPPSDQ 480
 DB 421 SSSDALRYKRPSTLLDGTNKLCOFSPVOELSEQPTSPDKEKASIRKLOTAPPSDQ 480
 QY 481 SKRLSVRTSSSGTQORSILSPILHRSQVEDNHYHTSPFLGISTSQOHLTKSAGLGLKGM 540
 DB 481 SKRLSVRTSSSGTQORSILSPILHRSQVEDNHYHTSPFLGISTSQOHLTKSAGLGLKGM 540
 QY 541 SDILAPQSTPSTLTSWYFATESSHFYGASAIYGGASAYSVCQPLTCGDOYVSVRR 600
 DB 541 SDILAPQSTPSTLTSWYFATESSHFYGASAIYGGASAYSVCQPLTCGDOYVSVRR 600
 QY 601 QKPSRPAASRSRMSHEESPFEPKQFKRRSCOMEGESIMSENSRELGKVGSSSTSGME 660
 DB 601 QKPSRPAASRSRMSHEESPFEPKQFKRRSCOMEGESIMSENSRELGKVGSSSTSGME 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665
 RESULT 2
 ABR52407
 ID ABR52407 standard; protein; 665 AA.
 XX
 AC ABR52407;
 XX
 DT 19-JUN-2003 (first entry)
 DE Protein relating to the invention SEQ ID NO: 148.
 XX
 KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;
 KW antipsoriatic; cardiant; cytostatic; gene therapy; liver disease;
 KW proliferative disorder; renal failure; cardiovascular disorder;
 KW immunological disorder; arthritis; psoriasis; congenital heart defect;
 KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
 XX
 OS Homo sapiens.
 XX
 PN WO200257460-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 20-DEC-2001; 2001WO-US050459.
 XX
 PF 20-DEC-2000; 2000US-0256868P.
 PR 30-MAR-2001; 2001US-0280186P.
 PR 01-MAY-2001; 2001US-0287735P.
 PR 05-JUN-2001; 2001US-0295848P.
 PR 25-JUN-2001; 2001US-0300465P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
 PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
 PI Krycek S, Mcatee P, Suchard S, Banas D;
 DR WPI; 2002-599721/64.
 DR N-PSDB; ACC60559.
 XX
 PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 PS Disclosure; Fig 19; 801pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid comprising a
 CC polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotides fully defined in the specification. The polynucleotide of
 CC the invention has antiproliferative, hepatotropic, nephrotropic,

CC antiarthritic, antipneumatic, cardiant, and cytostatic activity. The
 CC polynucleotide may have a use in gene therapy. A polynucleotide or
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention

XX
 XX Sequence 665 AA;

Query Match 100.0%; Score 665; DB 5; Length 665;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAHEMIGTIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRRLOO 60
 DB 1 MAHEMIGTIVTERLVALLSGTEKVLIDSRPFVEYNTSHLEAININCSKLMKRRLOO 60
 QY 61 DAVLTTELIOHSAKAKVIDCSQKVVYVYQSSQDVASLSSDCLTVLLGKLEKSNVHL 120
 DB 61 DAVLTTELIOHSAKAKVIDCSQKVVYVYQSSQDVASLSSDCLTVLLGKLEKSNVHL 120
 QY 121 LAGGFAEFSRCPFGICEGKSTLVPTCISQCLPVANIGPTRLIPNLVYLGQORDVUNKELI 180
 DB 121 LAGGFAEFSRCPFGICEGKSTLVPTCISQCLPVANIGPTRLIPNLVYLGQORDVUNKELI 180
 QY 181 QONGIGYVNASAYTCPPDPFIESHFLRPVNDSPCEKILPMLDKSVDETERAKASNGCV 240
 DB 181 QONGIGYVNASAYTCPPDPFIESHFLRPVNDSPCEKILPMLDKSVDETERAKASNGCV 240
 QY 241 LVHCLAGISRSATIAIYIMKMDNSLDEAYRFVKEKERTISPENFLQQLIDYEKKIN 300
 DB 241 LVHCLAGISRSATIAIYIMKMDNSLDEAYRFVKEKERTISPENFLQQLIDYEKKIN 300
 QY 301 QNGAGSPKSKLTLHLEKNEPVPVASEGGQSEPTLSPCCADSAEAAAGRPVHPASV 360
 DB 301 QNGAGSPKSKLTLHLEKNEPVPVASEGGQSEPTLSPCCADSAEAAAGRPVHPASV 360
 QY 361 PSVPSPVQSLBDSPLVQALSGHLASADRLSDSNKLRFSFLDIKVSYSASMAASIHGF 420
 DB 361 PSVPSPVQSLBDSPLVQALSGHLASADRLSDSNKLRFSFLDIKVSYSASMAASIHGF 420
 QY 421 SSSSDALEYKPTSTIDGTNKLCOFSPVQELSEQTPETSPDXEASIPKKLQTPARPSDQ 480
 DB 421 SSSSDALEYKPTSTIDGTNKLCOFSPVQELSEQTPETSPDXEASIPKKLQTPARPSDQ 480
 QY 481 SKRLHSVTRSSSGTARSLSLPLHRGSEVDNYHTSFLGLSTSOQHLTKSAGLKGWH 540
 DB 481 SKRLHSVTRSSSGTARSLSLPLHRGSEVDNYHTSFLGLSTSOQHLTKSAGLKGWH 540
 QY 541 SPILAPQSTPSTLSSWYFATESSHFYASATYGASYSAYSCSOLPTCGQOVYVRRR 600
 DB 541 SPILAPQSTPSTLSSWYFATESSHFYASATYGASYSAYSCSOLPTCGQOVYVRRR 600
 QY 601 QKPSRADRRSWMHESPFKQFKRRSCQMEFGESIMSENRRSEELGKVGQSQSFSGSME 660
 DB 601 QKPSRADRRSWMHESPFKQFKRRSCQMEFGESIMSENRRSEELGKVGQSQSFSGSME 660
 QY 661 IIEVS 665
 DB 661 IIEVS 665

RESULT 3

AAU79159 standard; protein; 517 AA.

AC AAU79159;

XX 02-JUL-2002 (first entry)

XX Human dual-specificity phosphatase-3 (DSP-16) alternative form protein.
 DE Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KW cancer; graft-versus-host disease; allergy; metabolic disease;
 KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KW intercellular adhesion; DSP-16 modulator; chromosome 12p.

OS Homo sapiens.

Key Location/Qualifiers

FT Domain 94..103

XX /label= Active_site_domain

PN WO200226997-A2.

PD 04-APR-2002.

PF 25-SEP-2001; 2001WO-US030124.

PR 26-SEP-2000; 2000US-0235487P.

PA (CEPT-) CEPTYR INC.

PI Lucite RM, Wei B;

DR WPI; 2002-315802/35.

XX N-PSDB; ABK48378.

PT New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.

XX Claim 50; Fig 4; 87pb; English.

XX The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity, for modulation of a proliferative response in a cell, survival
 CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present amino acid
 CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 CC alternative form protein of the invention. This sequence is encoded by
 CC the human DSP-16 gene located on chromosome 12p

SO Sequence 517 AA;

Query Match 71.0%; Score 472; DB 5; Length 517;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCPKDPFIESHFLRPVNDSPCEKILPMLDSVPTIEKAKASNGCVLVHCLAGISRSAT 253
 DB 46 TCPKDPFIESHFLRPVNDSPCEKILPMLDSVPTIEKAKASNGCVLVHCLAGISRSAT 105
 QY 254 IAIAYIMKMDNSLDEAYRFVKEKERTISPENFLQQLIDYEKKIKNQTGASGPKSKLTL 313
 DB 106 IAIAYIMKMDNSLDEAYRFVKEKERTISPENFLQQLIDYEKKIKNQTGASGPKSKLTL 165
 QY 314 LHEKXNEPVPVASEGGQSEPTLSPCCADSAEAAAGRPVHPASVPSVQSLBDS 373
 DB 166 LHEKXNEPVPVASEGGQSEPTLSPCCADSAEAAAGRPVHPASVPSVQSLBDS 225
 QY 374 SPLVQALSGHLASADRLSDSNKLRFSFLDIKVSYSASMAASLHGSSSDALEYKKS 433

KW liver disorder; erythroid associated disorder; haemolytic anaemia;
 KW cellular proliferative; differentiative disorder; leukaemia;
 KW acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer;
 KW immunogen.
 XX Homo sapiens.
 OS
 XX
 XX
 FH Key Location/Qualifiers
 FT Domain 11..131
 FT 158..297
 FT Domain /label= Rhodanese-like domain
 FT /label= Catalytic domain
 FT /note= "Dual specificity phosphatase catalytic domain"
 FT 242..254
 FT Active-site /label= Tyrosine-specific_protein_phosphatase_active_site
 FT
 PN WO200173059-A2.
 XX
 PD 04-OCT-2001.
 XX
 PD 23-MAR-2001; 2001WO-US009477.
 XX
 PR 24-MAR-2000; 2000US-0191858P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Meyers RA;
 XX
 DR WPI; 2001-611635/70.
 DR N-PSDB; AAS14639.
 XX
 PT New human dual specificity polypeptides and nucleic acids for diagnosis
 of disease and treatment of e.g. liver disorders.
 XX
 XX Claim 9; Fig 1; 143pp; English.
 XX
 CC The invention relates to two novel human dual specificity phosphatases
 designated 2117 and 38692, the nucleic acids encoding them (including
 fragments, allelic variants, their complements or nucleic acids that
 hybridise to them) and antibodies raised against the proteins. The
 CC antibody is useful for detecting the presence of the polypeptide, and the
 CC nucleic acid fragments are useful for detecting the presence of the
 CC nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
 CC antisense sequences) are useful for modulating the activity or expression
 of the polypeptide/nucleic acid, useful for the treatment of e.g.
 CC hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
 CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
 CC listed in the specification) liver disorders, erythroid associated
 CC disorders (e.g. haemolytic anaemia) cellular proliferative or
 CC differentiative disorders, leukaemias (e.g. acute myeloid leukaemia),
 CC carcinomas, sarcomas, metastatic cancers. Agents modulating 2117 and
 CC 38692 are also useful for modulating the proliferation, survival,
 CC migration or differentiation of a 38692 or 2117-expressing cell. The
 CC polypeptide and nucleic acids are useful for identifying modulating
 CC agents. The present sequence represents the dual specificity phosphatase
 CC 2117
 XX
 XX Sequence 665 AA;
 SQ
 Query Match 71.0%; Score 472; DB 4; Length 665;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 194 TCCKPDPFIPESHFLRVAVNDSPCEKILPMWDKSVDFIEKAKSNGCVLVHCLAGISRAT 253
 DB 194 TCCKPDPFIPESHFLRVAVNDSPCEKILPMWDKSVDFIEKAKSNGCVLVHCLAGISRAT 253
 QY 254 IAAIYIMKRMDSIDEXYRVKKEKRPISNPNFLQGLDYEKIKKQGTASGPKSKTKL 313
 DB 254 IAAIYIMKRMDSIDEXYRVKKEKRPISNPNFLQGLDYEKIKKQGTASGPKSKTKL 313
 QY 314 LHLKNEBPVPAVSEGQKSETPLSPCADSATSEAGORPVHPASVSPVQPSILSD 373

DB 314 LHLKNEBPVPAVSEGQKSETPLSPCADSATSEAGORPVHPASVSPVQPSILSD 373
 QY 374 SPLVQALSGLHLSAURLSDSNLKRFSFLDIKSVSYASMASLIGFSSSEDALEYKPS 433
 DB 374 SPLVQALSGLHLSAURLSDSNLKRFSFLDIKSVSYASMASLIGFSSSEDALEYKPS 433
 QY 434 TTLDTGNKLCQFSPVQELSEQTPETSPDKEASIRPKLQTAAPSDSQSKRLSVRTSSG 493
 DB 434 TTLDTGNKLCQFSPVQELSEQTPETSPDKEASIRPKLQTAAPSDSQSKRLSVRTSSG 493
 QY 494 TAQRSLSPHRSKSGVEDNVTSTFLGLSTSQOHLTKSAGLCKGMSDILAPQSTPBL 553
 DB 494 TAQRSLSPHRSKSGVEDNVTSTFLGLSTSQOHLTKSAGLCKGMSDILAPQSTPBL 553
 QY 554 TSSWYFATSSHFFYSASAIYGGASAYSCGOLPTCCGQVYSVRROKPSRADSRSBG 613
 DB 554 TSSWYFATSSHFFYSASAIYGGASAYSCGOLPTCCGQVYSVRROKPSRADSRSBG 613
 QY 614 HESPEKQFKRRSCOMERGESIMSENRREBELGKVSQSPFGSGMEITEVS 665
 DB 614 HESPEKQFKRRSCOMERGESIMSENRREBELGKVSQSPFGSGMEITEVS 665
 RESULT 6
 AAU79156
 ID AAU79156 standard; protein; 665 AA.
 AC AAU79156;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Human dual-specificity phosphatase-3 (DSP-16) protein.
 XX
 KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KW cancer; graft-versus-host disease; allergy; metabolic disease;
 KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KW intercellular adhesion; DSP-16 modulator; chromosome 12p.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 242..251
 FT /label= Active_site_domain
 PN WO200226997-A2.
 XX
 PD 04-APR-2002.
 XX
 PF 25-SEP-2001; 2001WO-US030124.
 XX
 PR 26-SEP-2000; 2000US-0235487P.
 XX
 XX (CEPT-) CEPTYR INC.
 PA
 PI Lucine RM, Wei B;
 XX
 DR WPI; 2002-315802/35.
 DR N-PSDB; ABK47596.
 XX
 PT New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 XX
 PS Claim 1; Fig 2; 87pp; English.
 XX
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity, for modulation of a proliferative response in a cell, survival

CC of a cell, or differentiation of a cell. The cell displays contact
CC inhibition of cell growth or anchorage independent growth and may display
CC altered intercellular adhesion. The agent may modulate apoptosis, or the
CC cell cycle. The identified modulators can be used to treat Duchenne
CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
CC cell proliferation, and cell cycle abnormalities. The present amino acid
CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
CC protein of the invention. This sequence is encoded by the human DSP-16
CC gene located on chromosome 12p
XX
XX
SQ Sequence 665 AA;

Query Match 71.0%; Score 472; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCPKPDFPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRAT 253
DB 194 TCPKPDFPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRAT 253
QY 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNFLGQLLDYEKKIKNOTGASGPKSKLT 313
DB 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNFLGQLLDYEKKIKNOTGASGPKSKLT 313
QY 314 IHLKPNBPVPAVSEGGOKSETPPLSPCADSATSEAAQRPVHPASVSPVQPSLLD 373
DB 314 IHLKPNBPVPAVSEGGOKSETPPLSPCADSATSEAAQRPVHPASVSPVQPSLLD 373
QY 374 SPLVQALSGHLASADRLSDSNKLRKSFSLDIKSVSYASMAASLHGFSSSEDALEYKRS 433
DB 374 SPLVQALSGHLASADRLSDSNKLRKSFSLDIKSVSYASMAASLHGFSSSEDALEYKRS 433
QY 434 TTLDTGNKLCQFSPVOELSEQTPETSPDKEASIPKKLTQAPSDSQSKRLHSVRTSSG 493
DB 434 TTLDTGNKLCQFSPVOELSEQTPETSPDKEASIPKKLTQAPSDSQSKRLHSVRTSSG 493
QY 494 TQORSLSLPLHRSGVEDNYHTSFLGLSTSOOHLTKSAGLKGWHSPIILAPQSTPSL 553
DB 494 TQORSLSLPLHRSGVEDNYHTSFLGLSTSOOHLTKSAGLKGWHSPIILAPQSTPSL 553
QY 554 TSSWYFATESSHFYASAIYGGASAYSCQLPTCGDQVYSVRRROKPSDRADSRSM 613
DB 554 TSSWYFATESSHFYASAIYGGASAYSCQLPTCGDQVYSVRRROKPSDRADSRSM 613
QY 614 HESPEKQFKRRSCQMEFGESIMSENRSREELGKVGSSQSFSGSMEIIEVS 665
DB 614 HESPEKQFKRRSCQMEFGESIMSENRSREELGKVGSSQSFSGSMEIIEVS 665

RESULT 7
AAU09946 standard; protein; 665 AA.
ID AAU09946;
AC AAU09946;
DT 18-JUN-2002 (first entry)
DE Protein sequence of human (dual specificity phosphatase) DSP-10.
XX
XX Human; dual specificity phosphatase; DSP-10; cancer; epilepsy; stroke;
XX neuronal degeneration syndrome; Alzheimer's disease; depression;
XX schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
XX osteoporosis; diabetes.
XX Homo sapiens.
XX OS
XX PN WO200177340-A1.
XX 18-OCT-2001.
XX PD 06-APR-2001; 2001WO-EP003966.
XX PF
XX

PR 10-APR-2000; 2000EP-00107143.
XX
XX (MERCK) MERCK PATENT GMBH.
XX
XX
XX
XX Duecker K;
XX
XX
XX MPI: 2002-010917/01.
XX
XX DR N-PsDB; AAS15768.
XX

PT Novel dual specificity phosphatase polypeptides useful for treating
PT cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
PT disease, depression, schizophrenia, asthma and immune disorders.
XX
XX

PS Claim 2; Page 37-39; 43pp; English.

CC The present invention relates to a new isolated dual specificity
CC phosphatase (DSP10) polypeptide, comprising a 665 residue amino acid
CC sequence that is fully defined in the specification. The invention also
CC provides a sequence encoded by a 3059 nucleotide sequence fully defined
CC in the specification, and a sequence having at least 95 % identity to the
CC polypeptide, or fragments or variants of DSP-10. The invention is useful
CC for treating cancer e.g. leukaemia, colon carcinoma, lung cancer,
CC prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
CC stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
CC schizophrenia, cardiac myopathies, asthma, immune disorders,
CC inflammatory processes e.g. arthritis, bowel disease, type I diabetes,
CC osteoporosis, diabetes and diabetes associated diseases. The molecules of
CC the invention are also useful as vaccines for inducing immunological
CC response in a mammal, in disease diagnosis and in assays for screening
CC agonistic or antagonistic compounds. Other uses of the invention include
CC identifying membrane bound or soluble receptors, as a diagnostic reagent,
CC in chromosome localisation studies, and as a valuable tool in tissue
CC expression studies. The present sequence represents the dual specificity
CC phosphatase, DSP-10, protein of the invention
XX
XX
SQ Sequence 665 AA;

Query Match 71.0%; Score 472; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCPKPDFPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRAT 253
DB 194 TCPKPDFPESHFLRPVNDSPCEKILPMLDKSVDFIEKAKASNGCVLVHCLAGISRAT 253
QY 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNFLGQLLDYEKKIKNOTGASGPKSKLT 313
DB 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNFLGQLLDYEKKIKNOTGASGPKSKLT 313
QY 314 IHLKPNBPVPAVSEGGOKSETPPLSPCADSATSEAAQRPVHPASVSPVQPSLLD 373
DB 314 IHLKPNBPVPAVSEGGOKSETPPLSPCADSATSEAAQRPVHPASVSPVQPSLLD 373
QY 374 SPLVQALSGHLASADRLSDSNKLRKSFSLDIKSVSYASMAASLHGFSSSEDALEYKRS 433
DB 374 SPLVQALSGHLASADRLSDSNKLRKSFSLDIKSVSYASMAASLHGFSSSEDALEYKRS 433
QY 434 TTLDTGNKLCQFSPVOELSEQTPETSPDKEASIPKKLTQAPSDSQSKRLHSVRTSSG 493
DB 434 TTLDTGNKLCQFSPVOELSEQTPETSPDKEASIPKKLTQAPSDSQSKRLHSVRTSSG 493
QY 494 TQORSLSLPLHRSGVEDNYHTSFLGLSTSOOHLTKSAGLKGWHSPIILAPQSTPSL 553
DB 494 TQORSLSLPLHRSGVEDNYHTSFLGLSTSOOHLTKSAGLKGWHSPIILAPQSTPSL 553
QY 554 TSSWYFATESSHFYASAIYGGASAYSCQLPTCGDQVYSVRRROKPSDRADSRSM 613
DB 554 TSSWYFATESSHFYASAIYGGASAYSCQLPTCGDQVYSVRRROKPSDRADSRSM 613
QY 614 HESPEKQFKRRSCQMEFGESIMSENRSREELGKVGSSQSFSGSMEIIEVS 665
DB 614 HESPEKQFKRRSCQMEFGESIMSENRSREELGKVGSSQSFSGSMEIIEVS 665

RESULT 8
AAU75789
ID AAU75789 standard; protein; 665 AA.
AC AAU75789;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human protein phosphatase 7 (PP7) protein sequence.
XX
XX Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
KM acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
KM Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
KM dementia; Parkinson's disease; developmental disorder; Down's syndrome;
KM cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
KM melanoma; myeloma sarcoma.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Domain 11..131
FT /label= Rhodanese_like_domain
FT 15..170
FT /note= "Dual specificity protein phosphatase"
FT 85..298
FT /note= "VH1-type dual specificity protein phosphatase"
FT Domain 158..297
FT /label= Catalytic domain DSPc
FT /note= "Dual specificity phosphatase"
FT 220..281
FT /note= "Tyrosine specific protein phosphatases active
FT site tyr_phosphatase.ptf"
FT 242..254
FT /note= "Tyrosine phosphatase"
FT 242..252
FT /note= "Tyrosine specific protein phosphatases"
PN MO200210363-A2.
XX
PD 07-FEB-2002.
XX
PF 26-JUL-2001; 2001MO-US023716.
XX
PR 28-JUL-2000; 2000US-0221679P.
PR 03-AUG-2000; 2000US-022372P.
PR 10-AUG-2000; 2000US-0224309P.
PR 18-AUG-2000; 2000US-0226728P.
PR 30-AUG-2000; 2000US-0229254P.
PR 08-SEP-2000; 2000US-0231366P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;
PI Stewart EA, Gandhi AR, Patterson C, Lee EA, Hafalia ADL, Lu DAM;
PI Tribouley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;
PI Walla NK, Kearney L;
XX
XX MPI; 2002-188735/24.
DR N-PSDB; ABK14474.
XX
XX
PT New protein phosphatases, useful for diagnosing, treating or preventing
PT immune system disorders (e.g. Crohn's disease), neurological disorders
PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
PT cancers).
XX
PS Claim 1; Page 107-108; 117pp; English.
XX
XX The present invention relates to a new polypeptide, a naturally occurring
CC amino acid sequence at least 95 % identical to it, a biologically active
CC fragment of it or an immunogenic fragment of it. The polypeptides,
CC polynucleotides, agonists and antagonists are useful for diagnosing,
CC treating or preventing disorders associated with aberrant expression of

CC protein phosphatases (PP), particularly immune system disorders e.g.
CC acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
CC asthma or Crohn's disease, neurological disorders e.g. epilepsy,
CC Huntington's disease, dementia or Parkinson's disease, developmental
CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
CC or sarcoma. The present amino acid sequence represents human protein
CC phosphatase 7 (PP7) which is one of several human protein phosphatases
CC (AAU75783-AAU75792) of the invention
XX
SQ Sequence 665 AA;
XX
Query Match 71.0%; Score 472; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4/2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 194 TCEKDPFIPESHFLRVVNDSECEKILPWLKDSVDEIERAKASNGCVLHCLAGISRSAT 253
DB 194 TCEKDPFIPESHFLRVVNDSECEKILPWLKDSVDEIERAKASNGCVLHCLAGISRSAT 253
XX
QY 254 IAIAYIMKMDMSLDEAYRFVYKERTTISPNFNLGQLLDYEKKIKNOTGASGPKSKLKL 313
DB 254 IAIAYIMKMDMSLDEAYRFVYKERTTISPNFNLGQLLDYEKKIKNOTGASGPKSKLKL 313
XX
QY 314 LHLEKNEVPVAVSEGGQKSETPPLSPCADSATSEAAAGRPVHPAVSPVSPVQPSLLEP 373
DB 314 LHLEKNEVPVAVSEGGQKSETPPLSPCADSATSEAAAGRPVHPAVSPVSPVQPSLLEP 373
XX
QY 374 SPLVQALSGHLHSAADRLSDSNLKLKSFSLDIKSVSASMAAHLGFSSEDLLEYKPS 433
DB 374 SPLVQALSGHLHSAADRLSDSNLKLKSFSLDIKSVSASMAAHLGFSSEDLLEYKPS 433
XX
QY 434 TTLDTGNKLCOFSPVOELSEOTPEISPDKEASIPKLOTAPPSDQSRLLSVRTSSSG 493
DB 434 TTLDTGNKLCOFSPVOELSEOTPEISPDKEASIPKLOTAPPSDQSRLLSVRTSSSG 493
XX
QY 494 TQRSLSLPLHRSQSVEDNYHTSFLGLSTSOQHLTKSAGLIGKWHSDILAPQSTPPL 553
DB 494 TQRSLSLPLHRSQSVEDNYHTSFLGLSTSOQHLTKSAGLIGKWHSDILAPQSTPPL 553
XX
QY 554 TSSWYFATESSHFYSASAIYGSASAYSQSLPTCGDQVYSVRRRQKPSDRADRRRW 613
DB 554 TSSWYFATESSHFYSASAIYGSASAYSQSLPTCGDQVYSVRRRQKPSDRADRRRW 613
XX
QY 614 HBSPPKQFKRRSCOMFEGESIMSENREBELGKVSQSSFGSGMEITIEVS 665
DB 614 HBSPPKQFKRRSCOMFEGESIMSENREBELGKVSQSSFGSGMEITIEVS 665
XX
RESULT 9
ABB97946
ID ABB97946 standard; protein; 665 AA.
XX
XX ABB97946;
XX
DT 06-SEP-2002 (first entry)
XX
DE Human protein sequence #13.
XX
XX Human; brain; tonsil; hippocampus; foetal brain; diagnose.
XX
XX Homo sapiens.
OS
XX
PN WO200252005-A1.
XX
PD 04-JUL-2002.
XX
PF 20-DEC-2001; 2001MO-JP011217.
XX
PR 22-DEC-2000; 2000JP-00389742.
XX
PA (KAZU-) KAZUSA DNA RES INST FOUND.
PA (CELE-) CELESTAR LEXICO-SCI LTD.

XX PI Ohara O, Nagase T, Nakajima D;
 XX DR MPI: 2002-500762/53.
 XX DR N-PSDB; ABN83966.
 PT Genes and their expression products cloned from human cDNA libraries for
 PT treatment and diagnosis of diseases associated with their expression.
 XX
 PS Claim 1(a); Page 112-116; 238pp; Japanese.
 CC The invention relates to DNA encoding polypeptides directly cloned from
 CC cDNA libraries originating in adult whole brain, human consil, human
 CC adult hippocampus and human foetal whole brain. Polypeptides and
 CC polynucleotides of the invention may be used in the investigation of
 CC differential expression of the DNA sequences in normal subjects and
 CC disease patients. They may also be used in the production of antibodies,
 CC oligonucleotide probes and DNA chips for diagnosis and identification of
 CC drugs for treatment of diseases with which the DNA sequences are
 CC associated. The sequences given in records ABB97934-ABB97964 represent
 CC human proteins of the invention
 CC
 SQ Sequence 665 AA;

Query Match 71.0%; Score 472; DB 5; Length 665;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCRPPIFIESHFLRPVNDSECEKILPMIDKSVDFIEKAKSNGCVLVHCLAGISRAT 253
 DB 194 TCRPPIFIESHFLRPVNDSECEKILPMIDKSVDFIEKAKSNGCVLVHCLAGISRAT 253
 QY 254 IIAIYIMKMDMSLDAYRFVKEKRPITISPNFNLQGLDYEEKIKNOTGASGPKSKKL 313
 DB 254 IIAIYIMKMDMSLDAYRFVKEKRPITISPNFNLQGLDYEEKIKNOTGASGPKSKKL 313
 QY 314 LHLKENEVPVAVSEGGQKSETPSPCADSATSEAAAGORPVHPASVPSVPSVPSLLED 373
 DB 314 LHLKENEVPVAVSEGGQKSETPSPCADSATSEAAAGORPVHPASVPSVPSVPSLLED 373
 QY 374 SPVQALSGHLASADPLEDSNKLKRSFLDIKSVSISASMAASLHGFSSEDALEYKKS 433
 DB 374 SPVQALSGHLASADPLEDSNKLKRSFLDIKSVSISASMAASLHGFSSEDALEYKKS 433
 QY 434 TILDGNTKLCOPSPVOELSEOTPETSPDKEBASIPKKLOTARPSDSQSRHLSVTSSSG 493
 DB 434 TILDGNTKLCOPSPVOELSEOTPETSPDKEBASIPKKLOTARPSDSQSRHLSVTSSSG 493
 QY 494 TAQRSLISPLHRSQVEDNYHTSFLFGLSTSOOHLTKSAGLKGWHSIDILAPOTSTPGL 553
 DB 494 TAQRSLISPLHRSQVEDNYHTSFLFGLSTSOOHLTKSAGLKGWHSIDILAPOTSTPGL 553
 QY 554 TSSWTFATSSHFYSASAIYGGASAYSANCSGLPTCGOQVYSVRRQKPSRADSRRW 613
 DB 554 TSSWTFATSSHFYSASAIYGGASAYSANCSGLPTCGOQVYSVRRQKPSRADSRRW 613
 QY 614 HESPEPEKQPKRSCOMEFGESIMSENREELGKYGSSQSSFGSWEIIEVS 665
 DB 614 HESPEPEKQPKRSCOMEFGESIMSENREELGKYGSSQSSFGSWEIIEVS 665

RESULT 10
 AAU79929
 ID AAU79929 standard; protein; 665 AA.
 AC AAU79929;
 XX 02-JUL-2002 (first entry)
 DT Human dual specificity phosphatase 21117 protein.
 DE Human dual specificity phosphatase 21117 protein.
 XX Human, dual specificity phosphatase 21117; erythroid-related disorder;
 KW haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;

KW erythrocytosis; liver-related disorder; cancer.
 XX Homo sapiens.
 OS
 XX
 FH Key
 XX Domain
 FT 11..131
 FT /label= Rhodanese_like_domain
 FT 21..24
 FT /note= "Casein kinase II phosphorylation site"
 FT 91..94
 FT /note= "Casein kinase II phosphorylation site"
 FT 158..297
 FT /note= "Dual specificity phosphatase catalytic domain"
 FT 214..217
 FT /note= "Casein kinase II phosphorylation site"
 FT 242..254
 FT /note= "Tyrosine specific protein phosphatase active site"
 FT 242..254
 FT /note= "C-X5-R motif"
 FT 266..269
 FT /note= "Casein kinase II phosphorylation site"
 FT 369..372
 FT /note= "Casein kinase II phosphorylation site"
 FT 421..424
 FT /note= "Casein kinase II phosphorylation site"
 FT 434..437
 FT /note= "Casein kinase II phosphorylation site"
 FT 458..461
 FT /note= "Casein kinase II phosphorylation site"
 FT 508..511
 FT /note= "Casein kinase II phosphorylation site"
 FT 589..592
 FT /note= "Casein kinase II phosphorylation site"
 FT 612..615
 FT /note= "Casein kinase II phosphorylation site"
 FT 617..620
 FT /note= "Casein kinase II phosphorylation site"
 FT 642..645
 FT /note= "Casein kinase II phosphorylation site"
 FT /note= "Casein kinase II phosphorylation site"
 XX US2002034807-A1.
 XX
 PN 21-MAR-2002.
 XX
 PD 23-MAR-2001; 2001US-00816494.
 XX
 PF 24-MAR-2000; 2000US-0191858P.
 XX
 PR (MEYE/) MEYERS R A.
 XX
 PA
 XX
 PI Meyers RA;
 XX
 DR MPI: 2002-351088/38.
 DR N-PSDB; ABK49402.
 XX
 PT New nucleic acids, designated 38692 and 21117, encoding dual specificity
 PT phosphatases for treating cell proliferation and differentiation
 PT disorders including hematopoietic and erythroid-related disorders and
 PT cancers.
 XX
 PS Claim 8; Fig 1; 76pp; English.
 XX
 CC The present invention relates to new nucleic acids designated 38692 and
 CC 21117 encoding dual specificity phosphatase family members. The nucleic
 CC acid, polypeptide encoded by it, and antibody specific for the
 CC polypeptide may be used to diagnose and treat haematopoietic-related
 CC disorders such as leukemias and autoimmune diseases, erythroid-related
 CC disorders such as anaemias and erythrocytosis, liver-related disorders,
 CC and cancers, particularly of the breast, colon, adipose, prostate and
 CC lung. The present amino acid sequence represents the human dual
 CC specificity phosphatase 21117 protein of the invention, as described
 CC above

XX Sequence 665 AA;
SQ

Query Match 71.0%; Score 472; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 194 TCPKPFIPESHFLRPVNDSPCEKILPWLDKSVDFIEKAKSNGCVLVHCLAGISRSAT 253
DB 194 TCPKPFIPESHFLRPVNDSPCEKILPWLDKSVDFIEKAKSNGCVLVHCLAGISRSAT 253
OY 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNFLGQLDYEKKIKNQTGASGPKSKLKL 313
DB 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNFLGQLDYEKKIKNQTGASGPKSKLKL 313
OY 314 LHLEKNEBPVAVSEGGQKSETPLSPPCADSATSEAGORPVHPASVPSVQPSLLED 373
DB 314 LHLEKNEBPVAVSEGGQKSETPLSPPCADSATSEAGORPVHPASVPSVQPSLLED 373
OY 374 SPLVQALSGHLISADRLSDNSNKLKRSFSLDIKSVSYASMAASLHGFSSEDLLEYKPS 433
DB 374 SPLVQALSGHLISADRLSDNSNKLKRSFSLDIKSVSYASMAASLHGFSSEDLLEYKPS 433
OY 434 TTLDTGNKLCQFSPVOELSEQTPETSPDKEASIPKKLQTARPSDSQSKRLHSVRTSSG 493
DB 434 TTLDTGNKLCQFSPVOELSEQTPETSPDKEASIPKKLQTARPSDSQSKRLHSVRTSSG 493
OY 494 TAOBSLSPLRSGSVEDNYHTSFLGLSTSQOHLTKSAGLKGWHSIDLAPQISTPSL 553
DB 494 TAOBSLSPLRSGSVEDNYHTSFLGLSTSQOHLTKSAGLKGWHSIDLAPQISTPSL 553
OY 554 TSSWYFATESSHFYASAIYGSASYSAYSCQLPTCGDQVYSVRRQKPSDRADRRSM 613
DB 554 TSSWYFATESSHFYASAIYGSASYSAYSCQLPTCGDQVYSVRRQKPSDRADRRSM 613
OY 614 HESPEKQFKRSCOMEFGESIMSENRSREBELGKVGSSSFGSGMEIIEVS 665
DB 614 HESPEKQFKRSCOMEFGESIMSENRSREBELGKVGSSSFGSGMEIIEVS 665
```

RESULT 11

ABB97291
ID ABB97291 standard; protein; 665 AA.

XX ABB97291;

XX 27-JUN-2002 (first entry)

XX Novel human protein SEQ ID NO: 559.

XX Human; antihaemic; vulnerary; antiinflammatory; immunomodulator;
XX antifertility; cerebroprotective; cytoskeletal; rheumatic; gene therapy;
XX neuroprotective; antiParkinsonian; protein therapy; EST;
XX expressed sequence tag.

XX Homo sapiens.

XX WO200222660-A2.

XX 21-MAR-2002.

XX 10-SEP-2001; 2001WO-US026015.

XX 11-SEP-2000; 2000US-00659671.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Weinman T, Drmanac RT;

XX WPI; 2002-292408/33.
XX N-PSDB; ABN32477.

PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.

XX Example 2; SEQ ID NO 559; 509pp; English.

XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to regulate
XX haematopoiesis e.g. to treat aplastic anemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a protein of the invention

XX Sequence 665 AA;

Query Match 71.0%; Score 472; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 194 TCPKPFIPESHFLRPVNDSPCEKILPWLDKSVDFIEKAKSNGCVLVHCLAGISRSAT 253
DB 194 TCPKPFIPESHFLRPVNDSPCEKILPWLDKSVDFIEKAKSNGCVLVHCLAGISRSAT 253
OY 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNFLGQLDYEKKIKNQTGASGPKSKLKL 313
DB 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNFLGQLDYEKKIKNQTGASGPKSKLKL 313
OY 314 LHLEKNEBPVAVSEGGQKSETPLSPPCADSATSEAGORPVHPASVPSVQPSLLED 373
DB 314 LHLEKNEBPVAVSEGGQKSETPLSPPCADSATSEAGORPVHPASVPSVQPSLLED 373
OY 374 SPLVQALSGHLISADRLSDNSNKLKRSFSLDIKSVSYASMAASLHGFSSEDLLEYKPS 433
DB 374 SPLVQALSGHLISADRLSDNSNKLKRSFSLDIKSVSYASMAASLHGFSSEDLLEYKPS 433
OY 434 TTLDTGNKLCQFSPVOELSEQTPETSPDKEASIPKKLQTARPSDSQSKRLHSVRTSSG 493
DB 434 TTLDTGNKLCQFSPVOELSEQTPETSPDKEASIPKKLQTARPSDSQSKRLHSVRTSSG 493
OY 494 TAOBSLSPLRSGSVEDNYHTSFLGLSTSQOHLTKSAGLKGWHSIDLAPQISTPSL 553
DB 494 TAOBSLSPLRSGSVEDNYHTSFLGLSTSQOHLTKSAGLKGWHSIDLAPQISTPSL 553
OY 554 TSSWYFATESSHFYASAIYGSASYSAYSCQLPTCGDQVYSVRRQKPSDRADRRSM 613
DB 554 TSSWYFATESSHFYASAIYGSASYSAYSCQLPTCGDQVYSVRRQKPSDRADRRSM 613
OY 614 HESPEKQFKRSCOMEFGESIMSENRSREBELGKVGSSSFGSGMEIIEVS 665
DB 614 HESPEKQFKRSCOMEFGESIMSENRSREBELGKVGSSSFGSGMEIIEVS 665
```

RESULT 12

ADA54744
ID ADA54744 standard; protein; 665 AA.

XX ADA54744;

XX 20-NOV-2003 (first entry)

XX Human protein, SEQ ID 2312.

XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nocotropic;
XX Gene Therapy; human; secretory protein; membrane proteins; cancer;
XX inflammatory disease; osteoporosis; neurological disease.

XX Homo sapiens.

XX EPI293569-A2.

PD 19-MAR-2003.
 XX
 PF 21-MAR-2002; 2002EP-00006586.
 XX
 PR 14-SEP-2001; 2001JP-00328381.
 XX 24-JAN-2002; 2002US-0350435P.
 XX
 PA (HELI-) HELIX RES INST.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
 XX
 DR WPI; 2003-395539/38.
 DR N-PSDB; ADA53105.
 XX
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 XX
 PS Claim 14; SEQ ID NO 2312; 205pp; English.
 XX
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.
 XX
 SQ Sequence 665 AA;

Query Match 71.0%; Score 472; DB 6; Length 665;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCEKPPFIFESHFLRPVNDSPCEKILPMLDKSVDFIEKAKSNGCVLVHCLAGISRSAT 253
 DB 194 TCEKPPFIFESHFLRPVNDSPCEKILPMLDKSVDFIEKAKSNGCVLVHCLAGISRSAT 253
 QY 254 IAIAYIMKMDMLDEAYRFVYKRRPTISNPNFLOQLDYEKKIKNOGASGPKSKKL 313
 DB 254 IAIAYIMKMDMLDEAYRFVYKRRPTISNPNFLOQLDYEKKIKNOGASGPKSKKL 313
 QY 314 IHIKENEVPAYSEGQKSETPLSPCADSATSEAGQRPVPASVSPVQPSLLLED 373
 DB 314 IHIKENEVPAYSEGQKSETPLSPCADSATSEAGQRPVPASVSPVQPSLLLED 373
 QY 374 SPLVQALSGHLTADRLSDSNKLRKSFSLDIKSVYSASMAASLHGFSSEDALEYKPS 433
 DB 374 SPLVQALSGHLTADRLSDSNKLRKSFSLDIKSVYSASMAASLHGFSSEDALEYKPS 433
 QY 434 TTLDGNTKLCQSPVQDELSTQETSPDKKEASIPKKLQTPARSDSQSKRLHVSRTSSSG 493
 DB 434 TTLDGNTKLCQSPVQDELSTQETSPDKKEASIPKKLQTPARSDSQSKRLHVSRTSSSG 493
 QY 494 TAQRSLSPILHRSGSVYDNDVHTFGLSTQOHLTKSAGLKGWMSDILAPQOTSTPSL 553
 DB 494 TAQRSLSPILHRSGSVYDNDVHTFGLSTQOHLTKSAGLKGWMSDILAPQOTSTPSL 553
 QY 554 TSSWYFATESSHFYASATYGSASYSAYSCSLPTCGDQVYVRRRKQKPSDRADRSRW 613
 DB 554 TSSWYFATESSHFYASATYGSASYSAYSCSLPTCGDQVYVRRRKQKPSDRADRSRW 613
 QY 614 HESSPFKQKRRSCQMEPESTIMSENRREELGKVGSSSFGSGSMIIIVS 665
 DB 614 HESSPFKQKRRSCQMEPESTIMSENRREELGKVGSSSFGSGSMIIIVS 665
 RESULT 13
 ID AAU79161 standard; protein: 665 AA.
 XX
 AC AAU79161;

XX 02-JUN-2002 (first entry)
 PT
 DE Human dual-specificity phosphatase-3 (DSP-16) mutant protein #1.
 XX
 KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KW cancer; graft-versus-host disease; allergy; metabolic disease;
 KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KW intercellular adhesion; DSP-16 modulator; mutant; mutein.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key location/Qualifiers
 FT Misc-difference 213
 FT /note= "wild-type Asp substituted by Ala"
 XX
 PN MO200226997-AZ.
 XX
 PD 04-APR-2002.
 XX
 PF 25-SEP-2001; 2001WO-US030124.
 XX
 PR 26-SEP-2000; 2000US-0235487P.
 XX
 PA (CEPT-) CEPTYR INC.
 XX
 PI Luche RM, Wei B;
 PI
 DR WPI; 2002-315802/35.
 XX
 PT New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 XX
 PS Claim 46; Page; 87pp; English.
 XX
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity, for modulation of a proliferative response in a cell, survival
 CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present amino acid
 CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 CC mutant protein #1. Note: This sequence is not shown in the specification
 CC but is derived from the wild-type human DSP-16 (AAU79156) protein given
 CC in figure 2 of the specification
 XX
 SQ Sequence 665 AA;

Query Match 68.0%; Score 452; DB 5; Length 665;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 SFCEKILPMLDKSVDFIEKAKSNGCVLVHCLAGISRSATIAIAYIMKMDMSLDEAYRF 273
 DB 214 SFCEKILPMLDKSVDFIEKAKSNGCVLVHCLAGISRSATIAIAYIMKMDMSLDEAYRF 273
 QY 274 VKEKRPITSPNPNFLOQLDYEKKIKNOGASGPKSKKLHLKENEVPVAVSSEGQKS 333
 DB 274 VKEKRPITSPNPNFLOQLDYEKKIKNOGASGPKSKKLHLKENEVPVAVSSEGQKS 333
 QY 334 ETPLSPCADSATSEAGQRPVPASVSPVQPSLLLEDSPVLQALSGHLTADRLSDSN 393
 DB 334 ETPLSPCADSATSEAGQRPVPASVSPVQPSLLLEDSPVLQALSGHLTADRLSDSN 393

Db 334 ETPFLSPDCADSATSEAGRPVHPASVSPVSPVPSLLEDSPLVQALSGHLSDRLLEDS 393
 QY 394 NKLKRSFSLDIKSVSASMAASLHGSSSEDALEYKPEPTLDGNNKLCQFSPVOELSE 453
 Db 394 NKLKRSFSLDIKSVSASMAASLHGSSSEDALEYKPEPTLDGNNKLCQFSPVOELSE 453
 QY 454 QPETSPPDKEBASIPKKLQTPARPSDSQSKRLHSVRTSSGTAQRSLSPLRSGVEDNY 513
 Db 454 QPETSPPDKEBASIPKKLQTPARPSDSQSKRLHSVRTSSGTAQRSLSPLRSGVEDNY 513
 QY 514 HTSFLFGLSTSQOHLTKSAGLGLKGMHSDTLAPQTSPSTSSWYPTATSSSHFYASAIY 573
 Db 514 HTSFLFGLSTSQOHLTKSAGLGLKGMHSDTLAPQTSPSTSSWYPTATSSSHFYASAIY 573
 QY 574 GGSASYSAYSCQLPTCGDQVSVRRRQKPSDRADRSRSHESPPKQPKRRSCOMERG 633
 Db 574 GGSASYSAYSCQLPTCGDQVSVRRRQKPSDRADRSRSHESPPKQPKRRSCOMERG 633
 QY 634 ESIMSENRSREELGKVGSGSSFGSGMEIIEVS 665
 Db 634 ESIMSENRSREELGKVGSGSSFGSGMEIIEVS 665
 RESULT 14
 ID AAU79162 standard; protein; 665 AA.
 XX AAU79162;
 DT 02-JUL-2002 (first entry)
 DE Human dual-specificity phosphatase-3 (DSP-16) mutant protein #2.
 XX Human dual-specificity phosphatase-3; DSP-16; autoimmune disease;
 KM mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
 KM cancer; graft-versus-host disease; allergy; metabolic disease;
 KM abnormal cell growth; abnormal cell proliferation; contact inhibition;
 KM cell cycle abnormality; anchorage independent cell growth; apoptosis;
 KM intercellular adhesion; DSP-16 modulator; mutant; mutein.
 XX Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 244 /note= "Wild-type Cys substituted by Ser"
 FT
 XX WO200226997-A2.
 PN 04-APR-2002.
 PD
 XX 25-SEP-2001; 2001WO-US030124.
 PF
 XX 26-SEP-2000; 2000US-0235487P.
 PR
 XX (CEPT-) CEPTYR INC.
 PA
 XX Luche RM, Wei B;
 PI
 XX
 DR WPI; 2002-315802/35.
 DT
 XX
 PT New DSP-16 polypeptide, useful for identifying modulators of its
 PT activity, which can be used in the treatment of disorders such as
 PT Duchenne muscular dystrophy, or cancer.
 XX
 PS Claim 46; Page; 87pp; English.
 XX
 CC The present invention relates to a new polypeptide, DSP-16, having a 665
 CC amino acid sequence, given in the specification, or a variant having at
 CC least 50 % identical residues, which retains the ability to
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The
 CC invention can be used for identifying agents which modulate DSP-16
 CC activity, for modulation of a proliferative response in a cell, survival

CC of a cell, or differentiation of a cell. The cell displays contact
 CC inhibition of cell growth or anchorage independent growth and may display
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the
 CC cell cycle. The identified modulators can be used to treat Duchenne
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
 CC cell proliferation, and cell cycle abnormalities. The present amino acid
 CC sequence represents the human dual-specificity phosphatase-3 (DSP-16)
 CC mutant protein #2. Note: This sequence is not shown in the specification
 CC but is derived from the wild-type human DSP-16 (AAU79156) protein given
 CC in figure 2 of the specification
 XX
 SQ Sequence 665 AA;
 Query Match 63.3%; Score 421; DB 5; Length 665;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 245 LAGISRSATITAIYIMKRMDSLDEAYRFVKEKRPISPNFNLGQLDYEKKIKNOTGA 304
 Db 245 LAGISRSATITAIYIMKRMDSLDEAYRFVKEKRPISPNFNLGQLDYEKKIKNOTGA 304
 QY 305 SGPKSXLKLHLKEKNEPVPAVSEGGQKSETPSPCADSATSEAGRPVHPASVSPVP 364
 Db 305 SGPKSXLKLHLKEKNEPVPAVSEGGQKSETPSPCADSATSEAGRPVHPASVSPVP 364
 QY 365 SVQPSLLEDSPLVQALSGHLSDRLLEDSNKLKRSFSLDIKSVSASMAASLHGSSSE 424
 Db 365 SVQPSLLEDSPLVQALSGHLSDRLLEDSNKLKRSFSLDIKSVSASMAASLHGSSSE 424
 QY 425 DALEYKPEPTLDGNNKLCQFSPVOELSEQTPETSPDKEBASIPKKLQTPARPSDSQSKL 484
 Db 425 DALEYKPEPTLDGNNKLCQFSPVOELSEQTPETSPDKEBASIPKKLQTPARPSDSQSKL 484
 QY 485 HSVRTSSGTAQRSLSPLRSGVEDNYHTSFLFGLSTSQOHLTKSAGLGLKGMHSDTL 544
 Db 485 HSVRTSSGTAQRSLSPLRSGVEDNYHTSFLFGLSTSQOHLTKSAGLGLKGMHSDTL 544
 QY 545 APQTSPLSTSSWYPTATSSSHFYASAIYGGSASYSAYSCQLPTCGDQVSVRRRQKPS 604
 Db 545 APQTSPLSTSSWYPTATSSSHFYASAIYGGSASYSAYSCQLPTCGDQVSVRRRQKPS 604
 QY 605 DRADRSRSHESPPKQPKRRSCOMERGESIMSENRSREELGKVGSGSSFGSGMEIIEV 664
 Db 605 DRADRSRSHESPPKQPKRRSCOMERGESIMSENRSREELGKVGSGSSFGSGMEIIEV 664
 QY 665 S 665
 Db 665 S 665
 RESULT 15
 ID ABR52352
 XX ABR52352 standard; protein; 665 AA.
 AC
 XX ABR52352;
 DT
 XX 19-JUN-2003 (first entry)
 DE
 XX Protein relating to the invention SEQ ID NO: 42.
 XX
 KM antiproliferative; hepatotropic; nephrotropic; antiarthritic;
 KM antipneumatic; cardiant; cyostatic; gene therapy; liver disease;
 KM proliferative disorder; renal failure; cardiovascular disorder;
 KM immunological disorder; arthritis; psoriasis; congenital heart defect;
 KM congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
 XX
 OS Homo sapiens.
 XX
 PN WO200257460-A2.
 XX
 PD 25-JUL-2002.
 XX

PF 20-DEC-2001; 2001WO-US050459.
XX
PR 20-DEC-2000; 2000US-0255686P.
PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
PI Krystek S, Mcatee P, Suchard S, Banas D;
XX
DR MPI: 2002-599721/64.
DR N-PSDB; ACC60521.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
PS Claim 5; Fig 5; 801pp; English.
XX
XX The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX
SQ Sequence 665 AA;
XX
Query Match 52.6%; Score 350; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 2e-317;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 316 LKPNRPVAVSEGGQKSETPSPPCADATSEAGQRPVHPASVSPVPSLLEDS 375
DB 316 LKPNRPVAVSEGGQKSETPSPPCADATSEAGQRPVHPASVSPVPSLLEDS 375
QY 376 LVQALSGHLASDRLEDSNKLKRSFSLDIKSVSYASMAASLHGFSSEDALEYKPTT 435
DB 376 LVQALSGHLASDRLEDSNKLKRSFSLDIKSVSYASMAASLHGFSSEDALEYKPTT 435
QY 436 LUGTNKLCQFSFVQELSEOTPETSPKXEPASIPKLTQTRPDSQSKRLHVRTSSGTA 495
DB 436 LUGTNKLCQFSFVQELSEOTPETSPKXEPASIPKLTQTRPDSQSKRLHVRTSSGTA 495
QY 496 QRSLSPLHRSQSVEDNHTSFLFGLSTSQOHLTKSAGLGLGWSHDILAPOTSPSLTS 555
DB 496 QRSLSPLHRSQSVEDNHTSFLFGLSTSQOHLTKSAGLGLGWSHDILAPOTSPSLTS 555
QY 556 SWYFATESHFSASAIYGSASYSAYSCSOLPTCGDYVSVRRQKPSDRADSRSSWHE 615
DB 556 SWYFATESHFSASAIYGSASYSAYSCSOLPTCGDYVSVRRQKPSDRADSRSSWHE 615
QY 616 BSEPFKQFRRSCQMEFGESIMSNNRRELGVGSSQSFSGSMETIEVS 665
DB 616 BSEPFKQFRRSCQMEFGESIMSNNRRELGVGSSQSFSGSMETIEVS 665
XX
RESULT 16
ABR52424
ID ABR52424 standard; protein; 664 AA.
XX

AC ABR52424;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein relating to the invention SEQ ID NO: 190.
XX
XX
XX antiproliferative; hepatotropic; nephrotropic; antiarthritic;
XX antipsoriatic; cardiant; cytostatic; gene therapy; liver disease;
XX proliferative disorder; renal failure; cardiovascular disorder;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
OS Homo sapiens.
XX
XX WO200257460-A2.
XX
PD 25-JUL-2002.
XX
XX 20-DEC-2001; 2001WO-US050459.
XX
XX 20-DEC-2000; 2000US-0255686P.
PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
PI Krystek S, Mcatee P, Suchard S, Banas D;
XX
DR MPI: 2002-599721/64.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
XX Example 57; Page 500-501; 801pp; English.
XX
XX The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX
SQ Sequence 664 AA;
XX
Query Match 42.3%; Score 281; DB 5; Length 664;
Best Local Similarity 100.0%; Pred. No. 4.9e-253;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 303 GASGPKSKLKLHLEKPNRPVAVSEGGQKSETPSPPCADATSEAGQRPVHPASVPS 362
DB 302 GASGPKSKLKLHLEKPNRPVAVSEGGQKSETPSPPCADATSEAGQRPVHPASVPS 361
QY 363 VPSVQPSLLEDSPLVQALSGHLASDRLEDSNKLKRSFSLDIKSVSYASMAASLHGFS 422
DB 362 VPSVQPSLLEDSPLVQALSGHLASDRLEDSNKLKRSFSLDIKSVSYASMAASLHGFS 421
QY 423 SEDALEYKPTTLDGNTKLCQFSFVQELSEOTPETSPKXEPASIPKLTQTRPDSQSK 482
DB 422 SEDALEYKPTTLDGNTKLCQFSFVQELSEOTPETSPKXEPASIPKLTQTRPDSQSK 481
XX

QY	483	RLHSVRTSSSGAORSLLEPLHRSSGVEDNHYHSPFLFGJLSTSOOHLTKSAGLGKGMHSD	542
Db	482	RLHSVRTSSSGAORSLLEPLHRSSGVEDNHYHSPFLFGJLSTSOOHLTKSAGLGKGMHSD	541
QY	543	ILAPOTSTPSLTSSWYFATESSHFYASAIYGGASAYSAYS	583
Db	542	ILAPOTSTPSLTSSWYFATESSHFYASAIYGGASAYSAYS	582
RESULT 17			
ID	AAB20325	standard; protein; 666 AA.	
XX	AAB20325;		
DT	29-MAY-2001	(first entry)	
XX			
DE	Human protein phosphatase and kinase protein-4.		
XX			
KW	Protein phosphatase and kinase protein; PPHK-4; human;		
KM	gastrointestinal disorder; immune system disorder; neurological disorder		
KW	cell proliferative disorder; cancer; diagnosis; therapy.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Modified-site	12	
FT		/note= "O-phosphorylated"	
FT	Modified-site	21	
FT		/note= "O-phosphorylated"	
FT	Modified-site	23	
FT		/note= "O-phosphorylated"	
FT	Modified-site	38	
FT		/note= "N-glycosylated"	
FT	Modified-site	38	
FT		/note= "N-glycosylated"	
FT	Modified-site	49	
FT		/note= "N-glycosylated"	
FT	Modified-site	72	
FT		/note= "O-phosphorylated"	
FT	Modified-site	82	
FT		/note= "O-phosphorylated"	
FT	Region	85..298	
FT		/note= "PH1-type dual specificity phosphatase signature"	
FT	Modified-site	91	
FT		/note= "O-phosphorylated"	
FT	Modified-site	190	
FT		/note= "N-glycosylated"	
FT	Modified-site	212	
FT		/note= "N-glycosylated"	
FT	Modified-site	214	
FT		/note= "O-phosphorylated"	
FT	Active-site	220..280	
FT		/note= "tyrosine specific protein phosphatase"	
FT	Region	237..278	
FT		/note= "y phosphatase signature"	
FT	Modified-site	266	
FT		/note= "O-phosphorylated"	
FT	Modified-site	280	
FT		/note= "O-phosphorylated"	
FT	Modified-site	300	
FT		/note= "N-glycosylated"	
FT	Modified-site	369	
FT		/note= "O-phosphorylated"	
FT	Modified-site	393	
FT		/note= "O-phosphorylated"	
FT	Modified-site	421	
FT		/note= "O-phosphorylated"	
FT	Modified-site	422	
FT		/note= "O-phosphorylated"	
FT	Modified-site	434	
FT		/note= "O-phosphorylated"	
FT	Modified-site	439	
FT		/note= "O-phosphorylated"	

FT	Modified-site	/note= "O-phosphorylated"
FT	468	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	471	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	479	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	528	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	590	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	597	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	605	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	610	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	613	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	618	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	628	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	641	/note= "O-phosphorylated"
FT	Modified-site	/note= "O-phosphorylated"
FT	643	/note= "O-phosphorylated"
PN	WO200120004-A2.	
PD	22-MAR-2001.	
PF	14-SEP-2000; 2000WO-US025515.	
PR	15-SEP-1999; 99US-0154141P.	
PA	(INCY-) INCYTE GENOMICS INC.	
PI	Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y,	
PI	Lu DAM;	
DR	WPI: 2001-244811/25.	
DR	N-PsDB; AAF30479.	
PT	Novel human protein phosphatase and kinase proteins for diagnosis,	
PT	treatment and prevention of gastrointestinal, immune system, neurological	
PT	and cell proliferative disorders.	
PS	Claim 1; Page 87-88; 103pp; English.	
XX	The present sequence is that of novel human protein phosphatase and	
XX	kinase protein PPHK-4, as predicted from Incyte Clone ID No. 1234795CB1	
CC	(see AAF30479). Tissues that express PPHK-4 (as a fraction of total	
CC	tissues expressing PPHK-4) include gastrointestinal (0.385),	
CC	cardiovascular (0.154), nervous (0.154) and reproductive (0.154).	
CC	Diseases or conditions associated with tissues expressing PPHK-4 (as a	
CC	fraction of total tissues expressing PPHK-4) include cancer (0.692),	
CC	inflammation or trauma (0.308) and cell proliferation (0.231). The	
CC	encoded protein shows homology to mouse neuronal tyrosine threonine	
CC	phosphatase 1. The invention provides human PPHK-1 to -11 polypeptides	
CC	(see AAB20322-32) and polynucleotides (see AAF30476-86). It also provides	
CC	expression vectors, host cells, antibodies, agonists and antagonists, as	
CC	well as methods for diagnosing, treating or preventing disorders	
CC	associated with expression of PPHK, including gastrointestinal	
CC	disorders, immune system disorders, neurological disorders and cell	
CC	proliferative disorders, including cancer	
XX	Sequence 666 AA;	

Query Match	40.0%;	Score 266;	DB 4;	Length 666;
Best Local Similarity	100.0%;	Pred. No. 4.9e-239;		
Matches 266;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY 194 TCPKPDFIPESHFLRPVNDSPFCCKILPWLKDSVPDIEKAKASNGCVLHCLAGISRAT 253
| | | | |
Db 194 TCPKPDFIPESHFLRPVNDSPFCCKILPWLKDSVPDIEKAKASNGCVLHCLAGISRAT 253
QY 254 IAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIKNOTGASGPKSKL 313
| | | | |
Db 254 IAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIKNOTGASGPKSKL 313
QY 314 LHLKRNKPNPVPVAVSEGGQKSTPLSPPCADSATSEAAQRPVHPASVPSVPSPSLLED 373
| | | | |
Db 314 LHLKRNKPNPVPVAVSEGGQKSTPLSPPCADSATSEAAQRPVHPASVPSVPSPSLLED 373
QY 374 SPLVQALSGHLHSADRLSDSNKLRSPSLDIKSVSYASMAASLHGFSSEDALEYKPS 433
| | | | |
Db 374 SPLVQALSGHLHSADRLSDSNKLRSPSLDIKSVSYASMAASLHGFSSEDALEYKPS 433
QY 434 TTLDTGNKLCOFSPVOELSEQTPEETS 459
| | | | |
Db 434 TTLDTGNKLCOFSPVOELSEQTPEETS 459

RESULT 18

AAM25744
ID AAM25744 standard; protein; 672 AA.

AC AAM25744;

DT 16-OCT-2001 (first entry)

DE Human protein sequence SEQ ID NO:1259.

KW Human; cancer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnery; antilicer; osteopathic; eczema;
KW dermatological; antiallergic; antisthmatic; antidiabetic; cytostatic;
KW immunoprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.

XX Homo sapiens.

XX OS

XX PN WO200153455-A2.

XX PD 26-JUL-2001.

XX PF 22-DEC-2000; 2000WO-US035017.

XX PR 23-DEC-1999; 99US-00471275.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX PA (HYSB-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI, 2001-457603/49.

XX DR N-PSDB; AAH99685.

XX PT Isolated human polynucleotides encoding polypeptides, useful for the
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX PS Claim 20; Page 260; 1217p; English.
XX CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
XX CC AAM25963. The proteins can have activities based on the tissues and cells

CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
CC antilicer; osteopathic; dermatological; antiallergic; antisthmatic;
CC antidiabetic; cytostatic; neuroprotective; immunoprotective; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders

XX SQ Sequence 672 AA;

Query Match 38.8%; Score 258; DB 4; Length 672;

Best Local Similarity 100.0%; Pred. No. 1.4e-211;

Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCPKPDFIPESHFLRPVNDSPFCCKILPWLKDSVPDIEKAKASNGCVLHCLAGISRAT 253
| | | | |
Db 201 TCPKPDFIPESHFLRPVNDSPFCCKILPWLKDSVPDIEKAKASNGCVLHCLAGISRAT 260
QY 254 IAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIKNOTGASGPKSKL 313
| | | | |
Db 254 IAIAYIMKMDMSLDEAYRFVKEKRPITSPNPNFLGQLLDYEKKIKNOTGASGPKSKL 320
QY 314 LHLKRNKPNPVPVAVSEGGQKSTPLSPPCADSATSEAAQRPVHPASVPSVPSPSLLED 373
| | | | |
Db 314 LHLKRNKPNPVPVAVSEGGQKSTPLSPPCADSATSEAAQRPVHPASVPSVPSPSLLED 380
QY 374 SPLVQALSGHLHSADRLSDSNKLRSPSLDIKSVSYASMAASLHGFSSEDALEYKPS 433
| | | | |
Db 374 SPLVQALSGHLHSADRLSDSNKLRSPSLDIKSVSYASMAASLHGFSSEDALEYKPS 440
QY 434 TTLDTGNKLCOFSPVOEL 451
| | | | |
Db 434 TTLDTGNKLCOFSPVOEL 458

RESULT 19

ABR52385
ID ABR52385 standard; protein; 660 AA.

XX ABR52385;

XX DT 19-JUN-2003 (first entry)

XX DE Protein relating to the invention SEQ ID NO: 114.

KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;
KW antiproliferative; cardiant; cytostatic; gene therapy; liver disease;
KW proliferative disorder; renal failure; cardiovascular disorder;
KW immunological disorder; arthritis; psoriasis; congenital heart defect;
KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.

XX Mus musculus.

XX OS

XX PN WO200257460-A2.

XX PD 25-JUL-2002.

XX PF 20-DEC-2001; 2001WO-US050459.

XX PR 20-DEC-2000; 2000US-0256868P.
XX PR 30-MAR-2001; 2001US-0280186P.

PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
PI Siemsek N, Bol D, Schieven G, Fingert J, Todderud CG, Bassolino D;
PI Kryszek S, Mcatee P, Suchard S, Banas D;
XX
DR WPI; 2002-599721/64.
DR N-PSDB; ACC60560.
XX
PT Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
PS Disclosure; Fig 12; 801pp; English.
XX
XX The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
SQ Sequence 660 AA;
XX
Query Match 9.6%; Score 64; DB 5; Length 660;
Best Local Similarity 100.0%; Pred. NO. 1.4e-50;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 114 SFNSVHLAAGFAFSGFPGTCEGKSTLVPTCISQPCLEPVANIGPRLIPNLVLCQORD 173
DB 114 SFNSVHLAAGFAFSGFPGTCEGKSTLVPTCISQPCLEPVANIGPRLIPNLVLCQORD 173
QY 174 VLNK 177
DB 174 VLNK 177
XX
RESULT 20
ADE08458
ID ADE08458 standard; protein; 579 AA.
XX
XX ADE08458;
AC
XX
DT 29-JUN-2004 (first entry)
XX
DE Novel protein (useful for identifying genetic disorders) #613.
XX
XX novel gene; novel protein; tissue marker; molecular weight marker;
KM chromosome marker; genetic disorder.
XX
XX Unidentified.
OS
XX
XX WO2003054152-A2.
PN
XX
PD 03-JUL-2003.
XX
PF 10-DEC-2002; 2002WO-US039555.
XX
PR 10-DEC-2001; 2001US-0339739P.
PR 11-DEC-2001; 2001US-0339453P.
PR 14-MAR-2002; 2002US-0365091P.

PR 14-MAR-2002; 2002US-0365384P.
PR 12-APR-2002; 2002US-0372381P.
PR 12-APR-2002; 2002US-0372615P.
PR 22-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.
XX
XX (HYGE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX
DR WPI; 2003-569235/53.
DR N-PSDB; ADE07547.
XX
XX New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX
PS Claim 20; SEQ ID NO 1524; 1177pp; English.
XX
XX The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present amino acid sequence represents a protein
CC of the invention.
SQ Sequence 579 AA;
XX
Query Match 3.0%; Score 20; DB 7; Length 579;
Best Local Similarity 100.0%; Pred. NO. 1.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 242 VHCAGISRSATIAIAYIMK 261
DB 186 VHCAGISRSATIAIAYIMK 205
XX
RESULT 21
ABR52382
ID ABR52382 standard; protein; 625 AA.
XX
XX ABR52382;
AC
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein relating to the invention SEQ ID NO: 110.
XX
XX antiproliferative; hepatotropic; nephrotropic; antiarthritic;
KM antipsoriatic; cardiant; cytostatic; gene therapy; liver disease;
KM proliferative disorder; renal failure; cardiovascular disorder;
KM immunological disorder; arthritis; psoriasis; congenital heart defect;
KM congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
XX Homo sapiens.
OS
XX
XX WO200257460-A2.
PN
XX
PD 25-JUL-2002.
XX
PF 20-DEC-2001; 2001WO-US050459.
XX
PR 20-DEC-2000; 2000US-0256868P.
PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX

PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
 PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Baasolino D;
 PI Krystek S, Mcatee P, Suchard S, Banas D;
 DR WPI; 2002-599721/64.
 XX
 PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 PS Disclosure; Fig 14; 801pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid comprising a
 CC polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotides fully defined in the specification. The polynucleotide of
 CC the invention has antiproliferative, hepatotropic, nephrotropic,
 CC antiarthritic, antipsoriatic, cardiant, and cyostatic activity. The
 CC polynucleotide may have a use in gene therapy. A polynucleotide or
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention
 SQ Sequence 625 AA;

Query Match 3.0%; Score 20; DB 5; Length 625;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 VHCLAGISRSATIAIAYIMK 261
 |||||
 DB 244 VHCLAGISRSATIAIAYIMK 263

RESULT 22
 ABR52350
 ID ABR52350 standard; protein; 625 AA.
 XX
 AC ABR52350;

DT 19-JUN-2003 (first entry)
 DE Protein relating to the invention SEQ ID NO: 39.

XX antiproliferative; hepatotropic; nephrotropic; antiarthritic;
 KW antipsoriatic; cardiant; cyostatic; gene therapy; liver disease;
 KW proliferative disorder; renal failure; cardiovascular disorder;
 KW immunological disorder; arthritis; psoriasis; congenital heart defect;
 KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
 XX

OS Homo sapiens.
 XX
 PN WO200257460-A2.
 PD 25-JUL-2002.

PF 20-DEC-2001; 2001WO-US050459.

PR 20-DEC-2000; 2000US-0256868P.
 PR 30-MAR-2001; 2001US-0280186P.
 PR 01-MAY-2001; 2001US-0287735P.
 PR 05-JUN-2001; 2001US-0295848P.
 PR 25-JUN-2001; 2001US-0300465P.

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
 PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Baasolino D;
 PI Krystek S, Mcatee P, Suchard S, Banas D;

XX
 DR WPI; 2002-599721/64.

PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 PS Disclosure; Fig 10; 801pp; English.

XX
 CC The invention relates to a novel isolated nucleic acid comprising a
 CC polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotides fully defined in the specification. The polynucleotide of
 CC the invention has antiproliferative, hepatotropic, nephrotropic,
 CC antiarthritic, antipsoriatic, cardiant, and cyostatic activity. The
 CC polynucleotide may have a use in gene therapy. A polynucleotide or
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention
 SQ Sequence 625 AA;

Query Match 3.0%; Score 20; DB 5; Length 625;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 VHCLAGISRSATIAIAYIMK 261
 |||||
 DB 244 VHCLAGISRSATIAIAYIMK 263

RESULT 23
 ABG73440
 ID ABG73440 standard; protein; 625 AA.
 XX
 AC ABG73440;

DT 10-MAY-2003 (first entry)

DE Human dual specific phosphatase 8 polypeptide.

XX Human; dual specific phosphatase 8; enzyme; infection; inflammation;
 KW tumour formation; cyostatic; antiinflammatory.
 XX

OS Homo sapiens.

PN US6482644-B1.

PD 19-NOV-2002.

PF 01-AUG-2001; 2001US-00920668.

PR 01-AUG-2001; 2001US-00920668.

PA (ISIS-) ISIS PHARM INC.

PI Cowseert LM;

DR WPI; 2003-298140/29.

DR N-PSDB; ABX10760.

PT New antisense compound targeted to a nucleic acid encoding human dual
 PT specific phosphatase 8, for modulating gene expression and treating
 PT diseases associated with expression of the phosphatase in humans.
 XX

PS Disclosure; Col 49-54; 36pp; English.

CC The invention relates to a compound targeted to the coding region of a
 CC nucleic acid encoding human dual specific phosphatase 8, where the

CC compound specifically hybridises with the region and inhibits the
CC expression of human dual specific phosphatase 8. The compound is useful
CC for inhibiting the expression of human dual specific phosphatase 8 in
CC cells or tissues, and for treating an animal, particularly a human,
CC suspected of having or being prone to a disease or condition associated
CC with expression of dual specific phosphatase 8. The compound is useful
CC for diagnosis, therapeutics and as a research reagent, e.g. to prevent
CC or delay infection, inflammation or tumour formation, and to distinguish
CC between functions of various members of a biological pathway. This
CC sequence represents human dual specific phosphatase 8

XX SQ Sequence 625 AA;

Query Match 3.0%; Score 20; DB 6; Length 625;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 VHCLAGISRSATTAIAYIMK 261
DB 244 VHCLAGISRSATTAIAYIMK 263

RESULT 24
ID AAM29150 standard; protein; 663 AA.

XX AAM29150;

DT 15-DEC-1997 (first entry)

XX Dual-specific murine threonine-tyrosine phosphatase M3/6.

KM murine; phosphatase; M3/6; dual specificity; Threonine; Tyrosine;
KM inactivate; mitogen activated protein kinase; MAP-K; cdc25 PTP; yeast;
KM tri-nucleotide repeat; repeat expansion; neurodegenerative; proliferative;
KM diagnosis; tumour; lung; brain; chromosomal deletion.

XX Mus sp.

PH Key Location/Qualifiers

FT Misc-difference 29..49 /note="region of homology with yeast cdc25"

FT Misc-difference 117..136 /note="region of homology with yeast cdc25"

FT Domain 244..253 /label=catalytic_domain

FT WO9706245-A1.

XX PD 20-FEB-1997.

XX PF 05-AUG-1996; 96WO-G8001906.

XX PR 04-AUG-1995; 95GB-00016059.

XX PA (MED1-) MEDICAL RES COUNCIL.

XX PI Davies KE, Theodorou A;

XX DR WPI; 1997-154253/14.

XX DR N-PSDB; AAT86758.

PT Murine mitogen activated protein (MAP) kinase phosphatase M3/6 - is a
PT suspected dual specificity Thr-Tyr phosphatase, useful for diagnosing and
PT treating neuro-degenerative or proliferative diseases e.g. tumours.

XX PS Claim 23; Fig 2; 51pp; English.

CC This is a murine phosphatase designated M3/6 (encoded by AAT86758). M3/6
CC is a suspected dual specificity Threonine-Tyrosine phosphatase, capable
CC of inactivating mitogen activated protein (MAP) kinase. The M3/6 protein
CC product shows high homology to the cdc25 PTP of yeast at residues 29-49
CC and 117-136. The gene also contains a complex triplet distal to the

CC catalytic domain which is translated into the protein. This domain
CC comprises a run of 4 serine residues which in turn is followed by a
CC further run comprising 23 serine residues which is interrupted near the N
CC terminal section by a single asparagine. This makes the phosphatase gene
CC a candidate for a human disease caused by repeat expansion or mutation.
CC M3/6 is expressed highly in the brain and may have utility in
CC investigating signal transduction mechanisms in brain and muscle. The
CC M3/6 and Hbs (a human homologue) genes may be responsible, if mutated,
CC for various neurodegenerative or proliferative diseases, and may
CC therefore be used for the diagnosis of such diseases, e.g. tumours,
CC especially lung or brain tumours, associated with deletion of the
CC chromosomal region 11p15.5. The polypeptides can be used to screen for
CC inhibitors to treat these diseases

XX SQ Sequence 663 AA;

Query Match 3.0%; Score 20; DB 2; Length 663;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 VHCLAGISRSATTAIAYIMK 261
DB 244 VHCLAGISRSATTAIAYIMK 263

RESULT 25

ID ABR52351 standard; protein; 663 AA.

XX ABR52351;

DT 19-JUN-2003 (first entry)

XX Protein relating to the invention SEQ ID NO: 40.

XX anti-proliferative; hepatotropic; nephrotropic; antiarthritic;

KM antipsoriatic; cardiac; cytosolic; gene therapy; liver disease;

KM proliferative disorder; renal failure; cardiovascular disorder;

KM immunological disorder; arthritis; psoriasis; congenital heart defect;

KM congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.

XX Mus musculus.

XX PN WO200257460-A2.

XX PD 25-JUL-2002.

XX PF 20-DEC-2001; 2001WO-US050459.

XX PR 20-DEC-2000; 2000US-0256868P.

XX PR 30-MAR-2001; 2001US-0280186P.

XX PR 01-MAY-2001; 2001US-0287735P.

XX PR 05-JUN-2001; 2001US-0295848P.

XX PR 25-JUN-2001; 2001US-0300465P.

XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;

XX PI Stiemers N, Bol D, Schieven G, Finger J, Todderud CG, Basolino D;

XX PI Krystek S, Mcatee P, Suchard S, Banas D;

XX DR WPI; 2002-599721/64.

PT Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.

XX PS Disclosure; Fig 10; 801pp; English.

CC The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,

CC antiarthritic, antipsoriatic, cardiac, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX
SQ Sequence 663 AA;

Query Match 3.0%; Score 20; DB 5; Length 663;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 242 VHCAGISRSATTAIAYIMK 261
Db 244 VHCAGISRSATTAIAYIMK 263

RESULT 26

ABB63527 standard; protein; 476 AA.

XX ABB63527;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 17373.

XX Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL07630.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

PS Disclosure; SEQ ID NO 17373; 21bp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLU16176-ABLU30511), expressed DNA
CC sequences (ABLU1840-ABLU16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 476 AA;

Query Match 2.0%; Score 13; DB 4; Length 476;

Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 246 AGISRSATTAIAY 258
Db 218 AGISRSATTAIAY 230

RESULT 27

ABG00724 standard; protein; 836 AA.

XX ABG00724;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #715.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS64911.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 20; SEQ ID NO 31083; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantifying a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (III) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 836 AA;

Query Match 1.8%; Score 12; DB 4; Length 836;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 VLWHCLAGTSRS 251
 |||||
 Db 744 VLWHCLAGTSRS 755

RESULT 28
 ADA44807
 ID ADA44807 standard; protein; 473 AA.
 XX
 AC ADA44807;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE CD4/TCR CD3epsilon chain chimeric protein CD4epsilon15, SEQ ID NO:2.
 XX
 KW HIV-1 infection; human immunodeficiency virus-1; CD4+ cell; chimeric CD4;
 KM endoplasmic reticulum; ER retention; envelope protein gp160;
 KM T cell receptor CD3epsilon chain; C-terminal domain; CD4epsilon15;
 KM gene therapy; human; receptor.
 XX
 OS Chimeric.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..458
 FT /label= CD4
 FT Region 459..473
 FT /note= "Part of the C-terminal domain of the T cell
 FT receptor CD3epsilon chain"
 XX
 PN MO2003076468-A1.
 XX
 PD 18-SEP-2003.
 XX
 PF 14-MAR-2003; 2003MO-ES000120.
 XX
 PR 14-MAR-2002; 2002ES-00000616.
 XX
 PA (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
 XX
 PI Alarcon Sanchez BJ, San Jose Martinez ME, Zaldivar Notario I;
 XX
 DR Gomez Buendia M,
 XX
 DR WPI; 2003-779059/73.
 XX
 DR N-PSDB; ADA44806.
 XX
 PT Composition for treating or preventing human immune deficiency virus,
 PT comprises CD4 chimeric protein having a protective effect in trans, or
 PT related nucleic acid.
 XX
 PS Claim 5; Page 33-35; 43pp; Spanish.
 XX
 CC The invention relates to a composition for the treatment or prevention of
 CC human immunodeficiency virus-1 (HIV-1) infection. The composition
 CC comprises CD4+ cells that have been transduced with a vector that encodes
 CC a chimeric CD4 molecule which is capable of being retained in the
 CC endoplasmic reticulum (ER). The invention also encompasses the use of a
 CC soluble protein factor produced by CD4+ cells that have been transduced
 CC with a vector encoding a chimeric CD4 protein; and the use of an
 CC expression system encoding a chimeric CD4 protein. The ER-localised
 CC chimeric CD4 molecule binds to the HIV-1 envelope protein gp160,
 CC resulting in HIV-1 retention in the ER and thereby preventing viral
 CC replication. In a specific embodiment, the chimeric CD4 molecule
 CC comprises CD4 fused to 15 amino acids of the C-terminal domain of the T
 CC cell receptor CD3epsilon chain; this chimeric CD4 molecule is designated
 CC CD4epsilon15 (ADA44807). A known chimeric CD4 of similar structure but
 CC containing only 10 amino acids from CD3epsilon can also be used.
 CC Compositions of the invention have an in trans effect on the replication
 CC of HIV-1, and may be used to treat and prevent HIV-1 infection. The
 CC present sequence represents the chimeric CD4 molecule CD4epsilon15, which
 CC is specifically claimed for use in compositions of the invention.

SO Sequence 473 AA;
 Query Match 1.2%; Score 8; DB 7; Length 473;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
 |||||
 Db 78 DRADSRRS 85

RESULT 29
 AAY59170
 ID AAY59170 standard; protein; 474 AA.
 XX
 AC AAY59170;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE CD4-Ig fusion protein CD4mmu.
 XX
 KW HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;
 KM secreted protein; SIV infection; medicament.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN CA1340741-C.
 XX
 PD 14-SEP-1999.
 XX
 PF 20-JUN-1989; 89CA-00588749.
 XX
 PR 20-JAN-1989; 89CA-00588749.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Seed B;
 XX
 DR WPI; 2000-063015/06.
 XX
 DR N-PSDB; AA248203.
 XX
 PT New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in
 PT the treatment of HIV or simian immunodeficiency virus infections.
 XX
 PS Example 1; Page 47-53; 89pp; English.
 XX
 CC The invention provides a fusion gene encoding a fusion protein that
 CC comprises an extracellular CD4 DNA sequence or its fragment which binds
 CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA
 CC sequence of an Ig heavy or light chain, where the DNA sequence encoding
 CC the variable region has been replaced with the DNA sequence which encodes
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is
 CC capable of being secreted. The fusion proteins are useful for treating
 CC HIV or SIV infections in animals, preferably humans. They are also useful
 CC for producing medicaments which can be used for treating HIV or SIV
 CC infections in humans. The present sequence represents the fusion protein
 CC CD4mmu where the CD4 is linked to human IgG1 at the Met2 site upstream of
 CC the CH1 region
 XX

SO Sequence 474 AA;
 Query Match 1.2%; Score 8; DB 3; Length 474;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
 |||||
 Db 78 DRADSRRS 85

RESULT 30
 AAP93011

ID	AAP93011 standard; protein; 481 AA.
XX	
AC	AAP93011;
XX	
D7	25-MAR-2003 (revised)
DT	03-AUG-1992 (first entry)
XX	
DE	Genetic construct which encodes CD4 linked to human Igm at the Pat site upstream of the CH2 region (fusion protein CD4Pmu).
DE	
KW	Fusion protein; immunoglobulin-like molecule; HIV; SIV; therapy; diagnosis; CD4; gp120; binding fragment; glycoprotein; variable region.
OS	Homo sapiens.
XX	
PX	EPJ25262-A.
PN	
PD	26-JUL-1989.
PF	
XX	20-JAN-1989; 89EP-00100913.
PR	
XX	22-JAN-1988; 88US-00147351.
PA	(GEHO) GEN HOSPITAL CORP.
PI	
Seed B;	
DR	WPI: 1989-214472/30.
DR	N-PsDB; AAN90359.
PT	Immunoglobulin-CD4 fusion proteins - used for treating HIV or SIV infections or detecting HIV or SIV in sample.
PS	
XX	Example; Table 4, Page 41-47; 68pp; English.
CC	The fusion protein genes of the invention pref. comprises cDNA sequences which encode CD4 or a fragment which binds gp120 ligated to an expression plasmid which encodes an antibody in which the variable region of the gene has been deleted (see WO87-02671). The CD4 portion of the fusion protein may comprise the complete CD4 sequence, the 370 AA extracellular region and the membrane spanning domain, or the extracellular region. CC Ig heavy chain is pref. from Igm, Igg1 or Igg3. The following are specifically claimed: fusion proteins CD4Hlambdai, CD4Mmu, CD4Pmu, CD4Etlambdai, and CD4Mmu (No. 67608), PCDAPlambda (No. 67609) and CC PCDAElambdai (No. 67610). (Updated on 25-MAR-2003 to correct PA field.) SQ Sequence 481 AA;
Query Match	1.2%; Score 8; DB 1; Length 481;
Best Local Similarity	100.0%; Pred. No. 1.9e+02;
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	605 DRADSRRS 612
DB	78 DRADSRRS 85
RESULT 31	
AAB19510	
ID	AAB19510 standard; protein; 481 AA.
XX	
AC	AAB19510;
XX	
DT	09-JAN-2001 (first entry)
XX	
DE	CD4-Igm fusion protein CH4Pmu.
XX	
KW	CD4; Igm; human; CD4Pmu; fusion protein; immunoglobulin; HIV; SIV; gp120; therapy; diagnosis.
OS	Homo sapiens.
Key	Location/Qualifiers
PH	

FT	Protein	1..395	
FT		/note="CD4 extracellular region"	
FT	Protein	400..481	
FT		/note="IgM heavy chain partial sequence"	
XX	US6117656-A.		
XX			
PD	12-SEP-2000.		
PF			
XX	07-JUN-1995;	95US-00479353.	
XX			
PR	22-JAN-1988;	88US-00147351.	
PR	23-JAN-1989;	89US-00239596.	
PR	09-JUN-1992;	92US-00896781.	
PR	12-APR-1993;	93US-00057952.	
PR	04-FEB-1994;	94US-00191708.	
XX			
PA	(GEHO) GEN HOSPITAL CORP.		
XX			
PI	Seed B;		
XX			
DR	WPI: 2000-586558/55.		
DR	N-PSDB; AAA50662.		
XX			
PT	CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or		
PT	SIV.		
XX			
PS	Example 1; Col 49-60; 39pp; English.		
XX			
CC	The present sequence is that of fusion protein CD4Pmu comprising the		
CC	extracellular portion of CD4, which binds to HIV gp120, linked at its C-		
CC	terminus to the human IgM heavy chain. To obtain the fusion protein, DNA		
CC	encoding CD4 was linked to IgM DNA at the Pst site upstream of the CH2		
CC	region (see AA50663). Fusion protein CD4Pmu and a nucleic acid encoding		
CC	it are claimed. Also claimed are a vector comprising the nucleic acid,		
CC	and a method of producing the fusion protein in secreted form using a		
CC	transformed host cell. The fusion protein may further comprise a		
CC	therapeutic agent, radiolabel or NMR imaging agent. The fusion protein		
CC	can be administered to an animal (including humans) for treatment of HIV		
CC	or SIV infection, and can also be used in assays for HIV or SIV, imaging		
CC	and tissue strains. IgM fusion proteins such as CD4Pmu provide complement-		
CC	mediated immunity		
XX			
SQ	Sequence 481 AA;		
Query Match	1.2%; Score 8; DB 3; Length 481;		
Best Local Similarity	100.0%; Pred. No. 1.9e+02;		
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	605 DRADSRSS 612		
Db	78 DRADSRSS 85		
RESULT 32			
AAVS1081			
ID	AAVS1081 standard; protein; 481 AA.		
XX			
AC	AAVS1081;		
XX			
DT	23-MAR-2000 (first entry)		
XX			
DE	Human fusion protein CD4Pmu.		
XX			
KW	Fusion protein; human; CD4; IgM; immunoglobulin; gp120;		
KW	anti-human immunodeficiency virus; CD4Pmu.		
XX			
OS	Homo sapiens.		
XX	Synthetic.		
XX			
PN	US6004781-A.		
XX			
DD	21-DEC-1999.		

```

XX PF 04-FEB-1994; 94US-00191708.
XX PR 22-JAN-1988; 88US-00147351.
XX PR 23-JAN-1989; 89US-00299596.
XX PR 09-JUN-1992; 92US-00896781.
XX PR 12-APR-1993; 93US-00057952.
XX PA (GEHO ) GEN HOSPITAL CORP.
XX PI Seed B;
XX DR WPI; 2000-085792/07.
XX DR N-PSDB; AA244064.
XX PT Fusion protein useful for the treatment of human immunodeficiency virus.
XX PS Example 1; Col 49-58; 39pp; English.
XX CC This invention describes a novel nucleic acid (I) encoding a fusion
XX CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
XX CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
XX CC chain (III). The products of the invention have anti-human
XX CC immunodeficiency virus (HIV) activity and are capable of binding to
XX CC gp120. The fusion protein is useful for treating human immunodeficiency
XX CC virus (HIV) or simian immunodeficiency virus (SIV). This sequence
XX CC represents the fusion protein CD4Fmu which is constructed from CD4 linked
XX CC to human IGM upstream of the CH2 region
XX SQ Sequence 481 AA;

Query Match 1.2%; Score 8; DB 3; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
DB 78 DRADSRSS 85

RESULT 33
AAVS9171
ID AAVS9171 standard; protein; 481 AA.
XX AC AAVS9171;
XX DT 14-MAR-2000 (first entry)
XX DE CD4-Ig fusion protein CD4Fmu.
XX KW HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;
XX KW secreted protein; SIV infection; medicament.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN CA1340741-C.
XX PD 14-SEP-1999.
XX PF 20-JAN-1989; 89CA-00588749.
XX PR 20-JAN-1989; 89CA-00588749.
XX PA (GEHO ) GEN HOSPITAL CORP.
XX PI Seed B;
XX DR WPI; 2000-063015/06.
XX DR N-PSDB; AA248204.
XX PT New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in
XX PT the treatment of HIV or simian immunodeficiency virus infections.

```

```

XX PS Example 1; Page 54-60; 89pp; English.
XX CC The invention provides a fusion gene encoding a fusion protein that
XX CC comprises an extracellular CD4 DNA sequence or its fragment which binds
XX CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA
XX CC sequence of an Ig heavy or light chain, where the DNA sequence encoding
XX CC the variable region has been replaced with the DNA sequence which encodes
XX CC extracellular CD4 or its gp120 binding fragment. The fusion protein is
XX CC capable of being secreted. The fusion proteins are useful for treating
XX CC HIV or SIV infections in animals, preferably humans. They are also useful
XX CC for producing medicaments which can be used for treating HIV or SIV
XX CC infections in humans. The present sequence represents the fusion protein
XX CC CD4Fmu where the CD4 is linked to human IgG1 at the Pst site upstream of
XX CC the CH2 region
XX SQ Sequence 481 AA;

Query Match 1.2%; Score 8; DB 3; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
DB 78 DRADSRSS 85

RESULT 34
AAB00158
ID AAB00158 standard; protein; 507 AA.
XX AC AAB00158;
XX DT 08-FEB-2001 (first entry)
XX DE sCD4-SCFv(17b) HIV single chain antibody fusion protein.
XX KW Fusion protein; HIV; human immunodeficiency virus; antibody; Fv; AIDS;
XX KW acquired immune deficiency syndrome; neutralisation; infection;
XX KW gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;
XX KW binding domain; single chain antibody; chimera; chimeric protein.
XX OS Human immunodeficiency virus.
XX OS Synthetic.
XX PN WO200055207-A1.
XX PD 21-SEP-2000.
XX PF 16-MAR-2000; 2000WO-US006946.
XX PR 16-MAR-1999; 99US-0124681P.
XX PA (USSH ) US NAT INST OF HEALTH.
XX PI Berger EA, Del Castillo CM;
XX DR WPI; 2000-638183/61.
XX DR N-PSDB; AA54045.
XX PT Novel neutralizing bispecific fusion proteins effective in viral such as
XX PT HIV neutralization, comprises two different binding domains, inducing-
XX PT binding domain and induced-binding domain functionally linked by linker.
XX PS Claim 39; Page 46-47; 55pp; English.
XX CC sCD4-SCFv(17b) is a neutralising bispecific fusion protein capable of
XX CC binding to two sites of its target protein. The protein comprises a first
XX CC binding domain capable of binding to an inducing site on the target
XX CC protein, a second binding domain capable of forming neutralising complex
XX CC with an induced epitope of the target protein and a linker connecting the
XX CC binding domains. sCD4-SCFv(17b) comprises a soluble CD4 fragment
XX CC (containing domains D1 and D2) fused to a single chain Fv portion of

```

CC antibody 17b via a linker. sCD4-SCFv(17b), its variant, analogue or
CC mimetic is used for inactivating gp120 protein of HIV, and for
CC neutralising HIV. It is also used for blocking and preventing the binding
CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte
CC CD4 and for inhibiting HIV replication. The chimeric proteins is
CC therefore useful for treating HIV infection and also AIDS. It is are
CC particularly useful in the prevention of infection during or immediately
CC after HIV exposure (e.g., mother/infant transmission, post-exposure
CC prophylaxis, and as a topical inhibitor) and for providing long term
CC resistance to HIV infections and AIDS. Gene therapy is used to secrete
CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal
CC or oral mucosa. The fusion proteins is highly potent, broadly cross-
CC reactive with neutralising antibody with high in vivo activity and no Fc-
CC mediated undesirable targeting properties. When the fusion protein is
CC substantially derived from human proteins, it has minimal immunogenicity
CC and toxicity in humans which is of great value in prevention of infection
CC during or immediately after HIV exposure

XX SQ Sequence 507 AA;

Query Match 1.2%; Score 8; DB 3; Length 507;

Best Local Similarity 100.0%; Pred. No. 2e+02; Mismatches 0; Gaps 0;

Matches 8; Conservative 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612

DB 78 DRADSRRS 85

RESULT 35

AAAR20152 standard; protein; 519 AA.

XX AAR20152;

DT 25-MAR-2003 (revised)

DT 31-MAR-1992 (first entry)

DE Human CD4 sequence encoded by PATY.6.

XX Human immunodeficiency virus; HIV; gp 120; AIDS; ARC; glycoprotein;

KM acquired immune deficiency syndrome; AIDS related complex;

XX T helper lymphocytes.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Peptide 1..25

FT /label= signal_sequence

XX WO9118618-A.

XX 12-DEC-1991.

PF 25-MAY-1990; 90US-00529186.

PR 25-MAY-1990; 90US-00529186.

XX (BIOJ) BIOGEN INC.

XX Fisher RA, Hession C, Burkly LC;

XX WPI; 1992-007200/01.

DR N-PSDB; AAQ20327.

XX New immuno-therapeutic human CD4 variants and derivs. - elicit AB

PT production to HIV gp.120, useful in treating, preventing and diagnosing

XX AIDS, ARC and HIV infections.

XX Disclosure; Fig 28; 179pp; English.

XX The sequence was deduced from the DNA sequence of subclone PATY.6, contg.

CC DNA coding for the full-length human CD4. The clone was constructed from

CC plasmids pBG178A and pBG378 (both in US8802940). The DNA can be used to
CC express recombinant CD4 and analogues for use in diagnosis and treatment
CC of diseases caused by infective agents whose primary targets are T4
CC lymphocytes. See also AAR20148-R20155 and AAR21078. (Updated on 25-MAR-
CC 2003 to correct PA field.)

XX SQ Sequence 519 AA;

Query Match 1.2%; Score 8; DB 2; Length 519;

Best Local Similarity 100.0%; Pred. No. 2.1e+02; Mismatches 0; Gaps 0;

Matches 8; Conservative 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612

DB 139 DRADSRRS 146

RESULT 36

AAAP94703 standard; protein; 524 AA.

XX AAP94703;

DT 25-MAR-2003 (revised)

DT 22-MAR-1991 (first entry)

DE Sequence encoded by T4 lymphocyte cDNA obtained from PBL clone lambda-203

XX -4.

XX HIV; soluble T4; immunotherapeutic; prophylactic; diagnostic; AIDS; ARC.

XX Homo sapiens.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Misc-difference 67 /note= "AA DESIGNATED NUMBER -23"

FT Misc-difference 90 /note= "AA DESIGNATED NUMBER 1"

FT Misc-difference 92 /note= "MATURE N-TERMINUS"

XX WO8901940-A.

XX 09-MAR-1989.

PF 01-SEP-1988; 88WO-US002940.

PR 04-SEP-1987; 87US-00094322.

XX 07-JAN-1988; 88US-00141649.

XX (BIOJ) BIOGEN INC.

XX Fisher RA, Gilbert W, Sato VL, Flavell RA, Maraganore JM;

XX WPI; 1989-085519/11.

DR N-PSDB; AAN90642.

XX DNA sequences coding for soluble T4-like polypeptide(s) - used in

PT immuno-therapeutic and immunosuppressive compns. and for preventing,

FT treating or detecting AIDS.

XX Disclosure; Fig 3; 207pp; English.

XX The polypeptides encoded are useful in immunotherapeutic, prophylactic

CC and diagnostic compns. They can be used to purify HIV from a sample. The

CC soluble T4 protein-based compns. are useful in treating immunodeficient

CC patients suffering from diseases caused by agents whose primary targets

CC are T4+ lymphocytes. They can be used for preventing, treating or

CC detecting AIDS, ARC and HIV infection. (Updated on 25-MAR-2003 to correct

CC PR field.)

XX Sequence 524 AA;

Query Match 1.2%; Score 8; DB 1; Length 524;

Best Local Similarity 100.0%; Pred. No. 2.1e+02; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612

Db 144 DRADSRSS 151

RESULT 37

AAR6783 AAR6783 standard; protein; 530 AA.

XX AAR6783;

AC 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 06-FEB-1993 (first entry)

XX CD4-IgG2 chimeric heavy chain.

XX homodimer; soluble CD4; T cell receptor; CD4 antigen; high recovery;

XX chimeric; increased serum half life; HIV infection; AIDS; ss.

OS Homo sapiens.

XX Chimeric.

XX Key

FT Location/Qualifiers

FT 1..205

FT /label= CD4 domain

FT 206..302

FT /label= CH1 domain

FT 303..312

FT /label= hinge domain

FT 313..423

FT /label= CH2 domain

FT 424..530

FT /label= CH3 domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

CD4-gamma-2 and CD4-IgG2 chimera(s) and expression vectors - for treatment, prevention and diagnosis of HIV infection.

Claim 15; Fig 4; 90pp; English.

This sequence represents a CD4-IgG2 chimeric heavy chain heterotrimer. It was produced by expression of the coding mutagenised cDNA (produced as described in AAQ28089) in Dhr-CHO cells. The protein is efficiently assembled intracellularly and effectively secreted from mammalian cells pref. CHO, COS, or myeloma cells as a heterotrimer, enabling high recovery and purification from the medium of cells expressing it. It possesses increased serum half-life and has increased avidity for HIV cf. heavy chain dimers. It can inhibit HIV infection of CD4+ cells and block the spread of HIV infection within a patient. Attachment to a detectable marker makes it useful in an assay for HIV or HIV infection, and it can also be linked to toxins, eg Diphtheria, Pseudomonas exotoxin A (domains I or II) or the deglycosylated A-chain of ricin. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS field)

SQ Sequence 530 AA;

Query Match 1.2%; Score 8; DB 2; Length 530;

Best Local Similarity 100.0%; Pred. No. 2.1e+02; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612

Db 78 DRADSRSS 85

RESULT 38

AAR6679 AAR6679 standard; protein; 530 AA.

XX AAR6679;

AC 25-MAR-2003 (revised)

DT 08-AUG-1994 (first entry)

XX CD4-IgG2 chimeric heavy chain.

XX CD4; gamma; heavy chain; chimeric; chimeric; immunconjugate; HIV;

XX human immunodeficiency virus; radionuclide; toxin; therapy; treatment;

XX imaging; detection; targeting; immunoglobulin; IgG.

OS Homo sapiens.

XX Key

FT Location/Qualifiers

FT 1..204

FT /label= CD4 Region.

FT 205..302

FT /label= CH1 Region.

FT 303..314

FT /label= Hinge Region.

FT 315..423

FT /label= CH2 Region.

FT 424..530

FT /label= CH3 Region.

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

Non-peptidyl toxin or radionuclide and CD4-gamma 2 or CD4-IgG2 immunoconjugates - used to kill HIV-infected cells and to image and stage HIV infection.

Disclosure; Fig 4; 142pp; English.

A tetramer comprising 2 IgG2 heavy chains or two CD4-IgG2 chimeric heavy chains and two kappa light chains or CD4-kappa light chains (AAR6680) linked to a non-peptidyl toxin or a gamma radiation-emitting radionuclide of low to moderate cytotoxicity. The resulting immunoconjugate comprising the toxin can be used to kill HIV infected cells and to treat HIV infected subjects to reduce the population of HIV infected cells. It can also be used to reduce the likelihood of infection. The immunoconjugate comprising the radionuclide can be used to image HIV infected tissue, to calculate the stage of HIV infection or the efficacy of an anti-HIV treatment using the imaging technique and for determining the prognosis of an HIV infected subject. (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 530 AA;
Query Match 1.2%; Score 8; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 605 DRADSRSS 612
|||||
DB 78 DRADSRSS 85
RESULT 39
AAV85080
ID AAV85080 standard; protein; 530 AA.
XX
XX AAV85080;
XX
XX 19-JUN-2000 (first entry)
XX
XX CD4-IgG2 chimeric heterotetramer heavy chain amino acid sequence.
DE
XX
XX CD4-IgG2 chimeric heavy chain heterotetramer; immunocjugate; treatment;
KM cytotoxic radionuclide; cell surface glycoprotein; prevent; infection;
KM cellular immune response interaction mediator; HIV interaction; staging;
KM prognosis; envelope glycoprotein burden; human.
XX
XX Homo sapiens.
XX
XX US6034223-A.
XX
XX 07-MAR-2000.
XX
XX 07-JUN-1995; 95US-00477460.
XX
XX 07-AUG-1992; 92US-00927931.
PR 06-AUG-1993; 93WO-US007422.
PR 03-FEB-1995; 95US-00379516.
XX
XX (PROG-) PROGENICS PHARM INC.
XX
XX Allaway GP, Maddon PJ;
XX
XX WPI; 2000-269502/23.
DR N-PSDB; AAZ98856.
XX
XX New immunocjugate, used to treat, prevent or image human immune
PT deficiency virus infection, comprises radionuclide attached to
PT heterotetramer of CD4-immunoglobulin chimerae.
XX
XX Disclosure; Fig 4; 58pp; English.
XX
XX This sequence represents the CD4-IgG2 chimeric heavy chain amino acid
CC sequence from the CD4-IgG2 chimeric heterotetramer. The invention relates
CC to an immunocjugate comprising a cytotoxic radionuclide and a
CC heterotetramer of two heavy chains and two light chains. The cytotoxic
CC radionuclide is linked to either the heavy chains or the light chains, or
CC to all four chains, directly or through a bifunctional chelator. Both
CC heavy chains are chimeric CD4-Ig (immunoglobulin) G2 chains encoded by
CC vector CD4-IgG2HC-PRCCMV (ATCC 75193) and both light chains are chimeric
CC CD4-kappa chains encoded by vector CD4-KLC-PRCCMV (ATCC 75194). CD4 is a
CC non-polymorphic cell surface glycoprotein that is expressed on the
CC surface of helper T lymphocytes, cells of the monocyte/macrophage lineage
CC and dendritic cells. CD4 associates with major histocompatibility complex
CC (MHC) class II molecules on the surface of antigen presenting cells to
CC mediate efficient cellular immune response interactions. In humans CD4 is
CC the target of interaction with the human immunodeficiency virus HIV. The
CC immunocjugate is used to kill cells infected with HIV, and for treating
CC or preventing infection. It is also used for imaging HIV-infected tissues
CC (for staging or prognosis of infection, and for assessing efficacy of
CC treatments). The immunocjugate is also used to determine the HIV
CC envelope glycoprotein burden, once determined, this information is used
CC in the staging and prognosis of HIV infected patients. The
CC immunocjugate should be active against all strains of HIV (since the

CC CD4-gp120 interaction is essential for infection). The heterotetramers
CC are assembled intracellularly and secreted efficiently from mammalian
CC cells, allowing high recovery and purification from the culture medium.
CC They have longer half-life in serum and greater avidity than heavy chain
CC dimers
XX
XX SQ Sequence 530 AA;
Query Match 1.2%; Score 8; DB 3; Length 530;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 605 DRADSRSS 612
|||||
DB 78 DRADSRSS 85
RESULT 40
AAB67323
ID AAB67323 standard; protein; 530 AA.
XX
XX AAB67323;
XX
XX 23-APR-2001 (first entry)
XX
XX CD4-IgG2 chimeric heavy chain protein.
DE
XX
XX Immunocjugate; chelator; chimeric; HIV; human immunodeficiency virus.
KM
XX Homo sapiens.
XX
XX US6177549-B1.
XX
XX 23-JAN-2001.
XX
XX 10-JUN-1999; 99US-00329916.
XX
XX 07-AUG-1992; 92US-00927931.
PR 06-AUG-1993; 93WO-US007422.
PR 03-FEB-1995; 95US-00379516.
PR 07-JUN-1995; 95US-00477460.
XX
XX (PROG-) PROGENICS PHARM INC.
XX
XX Maddon PJ, Allaway GP;
XX
XX WPI; 2001-158582/16.
DR
XX
XX Immunocjugate for treating human immunodeficiency virus-infected
PT subject, consists of cytotoxic radionuclide linked to heterotetramer
PT comprising two chimeric CD4-IgG2 heavy chains and two chimeric CD4-
PT kappa/light chains.
XX
XX Disclosure; Fig 4; 43pp; English.
XX
XX The present invention relates to an immunocjugate, comprising a
CC cytotoxic radionuclide linked, directly or via a bifunctional chelator,
CC to a heterotetramer having two chimeric CD4-IgG2 heavy chains encoded by
CC an expression vector CD4-IgG2HC-PRCCMV and two chimeric CD4-kappa light
CC chains encoded by an expression vector CD4-KLC-PRCCMV. The invention is
CC useful for killing human immunodeficiency virus (HIV)-infected cells, for
CC the treatment and prevention of infection with HIV
XX
XX SQ Sequence 530 AA;
Query Match 1.2%; Score 8; DB 4; Length 530;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 605 DRADSRSS 612
|||||
DB 78 DRADSRSS 85

DE	CD4-immunoglobulin G2 (IgG2) chimeric heterotetramer.
XX	
KM	CD4; immunoglobulin G2; Ig gamma2; human immunodeficiency virus-1; HIV-1;
XX	mutant; mutein.
OS	Homo sapiens.
XX	Synthetic.
FT	Key
FT	Peptide
FT	Protein
FT	Location/Qualifiers 1..25 /label= Signal_peptide 26..530 /note= "Mature CD4-IgG2 chimeric heterotetramer"
XX	
PN	US6451313-B1.
PD	17-SEP-2002.
XX	
PP	07-JUN-1995; 95US-00484681.
XX	
PR	08-FEB-1991; 91US-00653684.
PR	10-FEB-1992; 92WO-US001143.
PR	08-DEC-1992; 92US-00960440.
XX	
PA	(PROG-) PROGENICS PHARM INC.
XX	
P1	Maddon PJ, Beaudry GA;
DR	WPI; 2003-038273/03.
DR	N-Psdb; ABS55721.
XX	
PT	Novel CD4-immunoglobulin G2 chimeric heterotetramer neutralizes human
PT	immunodeficiency virus-1 with two heavy and light chains encoded by
PT	expression vectors designated CD4-IgG2HC-prcCMV and CD4-KLC-prcCMV,
PT	respectively.
XX	
PS	Claim 1; Fig 4A-H; 54pp; English.
XX	
CC	The invention describes a purified CD4-immunoglobulin (Ig)G2 chimeric
CC	heterotetramer (I) that neutralises human immunodeficiency virus-1 (HIV-
CC	1) having two heavy chains encoded by an expression vector designated CD4
CC	-IG2HC-prcCMV, and two light chains encoded by expression vector (II) or
CC	designated CD4-KLC-prcCMV. (I) and a composition (II) comprising (I) or
CC	(II) linked to a toxin, are useful for inhibiting HIV infection of a CD4
CC	cell, and preventing a subject being infected with HIV by blocking the
CC	spread of HIV infection. This is the amino acid sequence of the CD4-
CC	immunoglobulin G2 chimeric heavy chain chimeric heterotetramer useful in
CC	inhibiting HIV infection
XX	
SQ	Sequence 530 AA;
XX	
Query March	1.2%; Score 8; DB 6; Length 530;
Best local Similarity	100.0%; Pred. No. 2.1e+02;
Matches 8; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	605 DRADSRSS 612
Db	78 DRADSRSS 85
XX	
RESULT 43	
ID	AAR27278
XX	
AC	AAR27278; standard; protein; 532 AA.
XX	
DT	25-MAR-2003 (revised)
DT	28-JUL-1995 (first entry)
XX	
DE	CD4:gamma peptide chimeric protein.
XX	
KW	Fusion protein; CD4; extracellular domain; zeta; eta; gamma;
KW	membrane spanning domain; intracellular domain; type I;

```

KM integral membrane homodimer; TCR; T cell antigen receptor;
KW extracellular domain; mouse; human; receptor; chimera;
KW HPB-ALu tumour cell line; natural killer cell.
XX
OS Homo sapiens.
XX
PN WO9215322-A1.
XX
PD 17-SEP-1992.
XX
PF 06-MAR-1992; 92WO-US001785.
XX
PR 07-MAR-1991; 91US-00665961.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Seed B, Romeo C, Kolanus W;
XX
DR WPI; 1992-331474/40.
DR N-PSDB; AAQ28706.
XX
PT Therapeutic cells expressing chimeric receptors - directing cellular
PT response to an infective agent, useful in treating HIV-1, AIDS
PT Pneumocystis carinii infections etc.
XX
PS Example 2; Page 74-76; 114pp; English.
XX
CC This sequence represents a fusion protein between the CD4 extracellular
CC domain and the gamma protein membrane spanning domain and intracellular
CC domain. The Fc-receptor-associated gamma chain is expressed in cell
CC surface complexes with additional polypeptides, some of which mediate
CC ligand recognition, and others which have undefined function. Gamma bears
CC a homodimeric structure and overall organisation very similar to that of
CC zeta (see also AAQ28704), and is a component of both the mast
CC cell/basophil high affinity Ige receptor, Fc-epsilon-RI, which consists
CC of at least three distinct polypeptide chains and one of the low affinity
CC receptors for IgG, represented in mice by Fc-gamma-RII-alpha. In the
CC production of the CD4 receptor chimera, the gamma cDNA was isolated from
CC the HPB-ALu tumour cell line and from human natural killer cells. The
CC gamma cDNA was joined to the extracellular domain by engineering a BamHI
CC site just upstream of the membrane spanning domain, by a BamHI site
CC naturally present a few residues upstream of the membrane spanning
CC domain. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 532 AA;

Query Match          1.2%; Score 8; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 605 DRADSRRS 612
   |||||
Db 78 DRADSRRS 85

RESULT 44
ID AAR78678 standard; protein; 532 AA.
AC AAR78678;
XX
DE 16-APR-1996 (first entry)
XX
DE T-cell receptor eta.
XX
KW Chimeric receptor; CD4; T-cell receptor eta; HIV; cytolysis;
KW human immunodeficiency virus; adoptive immunotherapy.
XX
OS Homo sapiens.
XX
PN WO9521528-A1.
XX
PD 17-AUG-1995.

```

```

XX
PF 12-JAN-1995; 95WO-US000454.
XX
PR 14-FEB-1994; 94US-00195395.
PR 02-AUG-1994; 94US-00284391.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Seed B, Banapour B, Romeo C, Kolanus W;
XX
DR WPI; 1995-292893/38.
DR N-PSDB; AAQ96124.
XX
PT Target cytolysis of HIV-infected cells - by chimeric CD4 receptor-bearing
PT cells.
XX
PS Example 2; Page 78-79; 118pp; English.
XX
CC Fusion proteins comprising the extracellular domain of CD4 fused to T-
CC cell receptor zeta, gamma or eta (AAR78676-78, respectively) were
CC expressed in CV1 using a vaccine virus vector. These CD4:zeta, CD4:gamma
CC and CD4:eta chimeric receptors mediated cytolysis of targets expressing
CC HIV gp120/41
XX
SQ Sequence 532 AA;

Query Match          1.2%; Score 8; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 605 DRADSRRS 612
   |||||
Db 78 DRADSRRS 85

RESULT 45
ID AAR89458 standard; protein; 532 AA.
AC AAR89458;
XX
DE 26-SEP-1996 (first entry)
XX
DE CD4:eta fusion protein.
XX
KW CD7; transmembrane domain; chimeric receptor; CD5; CD34; CH2; CH3; IgG1;
KW human; CD4; HIV; proteinaceous alpha-helix; T cell; B cell; neutrophil;
KW dendritic cell; therapy; mammal; infection.
XX
OS Synthetic.
XX
PN WO9603883-A1.
XX
PD 15-FEB-1996.
XX
PF 26-JUL-1995; 95WO-US009468.
XX
PR 02-AUG-1994; 94US-00284391.
PR 24-FEB-1995; 95US-00394388.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Seed B, Banapour B, Romeo C, Kolanus W;
XX
DR WPI; 1996-129034/13.
DR N-PSDB; AAT10803.
XX
PT Membrane-bound chimeric receptor comprising extracellular portion
PT including CD4 fragment - cells expressing receptor can be used for
PT treatment of HIV infection.
XX
PS Example 2; Page 80-81; 134pp; English.

```

CC AAT10801-T10803 represent membrane bound proteinaceous chimeric receptors
 CC of the invention. This sequence represents the CD4:eta chimera. The
 CC transmembrane region of the chimeric receptor acts to separate the
 CC intracellular and extracellular domains of the chimera, and contains a
 CC portion of the CD7 (see AAR89440), CD5 or CD4 transmembrane domains.
 CC Alternatively, the extracellular portion of the receptor can be separated
 CC from the intracellular domain by the hinge, CH2 and CH3 domains of human
 CC IgG1 (see AAR89441). The extracellular portion of the chimeric receptor
 CC contains a fragment of CD4 (amino acids 1-394 or 1-200 of the CD4
 CC sequence, see AAR89450 and AAR89451) which specifically recognises and
 CC binds HIV-infected cells, but does not mediate HIV infection. The
 CC extracellular domain of the receptor is separated from the cell membrane
 CC by 48 or 72 angstroms, or by one or more proteinaceous alpha-helices. The
 CC cells expressing the receptor are preferably T cells, B cells,
 CC neutrophils, or dendritic cells. The therapeutic cells expressing the
 CC chimeric receptor are administered to a mammal to treat HIV infection

XX Sequence 532 AA;
 SQ

QY Query Match 1.2%; Score 8; DB 2; Length 532;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
 |||||
 78 DRADSRRS 85

Db

RESULT 46
 AAW02215
 ID AAW02215 standard; protein; 532 AA.
 XX
 AC AAW02215;
 XX
 DT 16-OCT-2003 (revised)
 DT 11-NOV-1996 (first entry)
 XX
 DE CD4:T-cell receptor eta chain chimeric receptor.
 XX
 KW Chimeric receptor; cellular immunity; adoptive immunotherapy; CD4;
 KW human immunodeficiency virus type 1; HIV-1; AIDS; therapy;
 KW T-cell receptor eta chain; cytotoxic T lymphocyte; CTL.
 XX
 OS Homo: sapiens.
 OS Mus sp.
 OS Chimeric.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..393
 FT /label= "Extracellular domain"
 FT /note= "CD4 extracellular domain"
 FT 394..396
 FT /label= Linker
 FT /note= "encoding DNA contains a BamHI site used for
 fusion construction"
 FT 397..532
 FT /note= "region of fusion derived from eta chain,
 preferred signal-transducing portions for constructs of
 the invention are amino acids 421-532, 423-455, 438-455,
 461-494, 494-528 or 400-420"
 FT 400..437
 FT /label= "Transmembrane domain"
 FT /note= "eta chain transmembrane domain"
 FT 438..575
 FT /label= "Intracellular domain"
 FT /note= "eta chain intracellular domain"
 XX
 PN W09625953-A1.
 XX
 PD 29-AUG-1996.
 XX
 PF 25-JAN-1996; 96WO-US001056.
 XX

PR 24-FEB-1995; 95US-00394176.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Seed B, Romeo C, Kolanus W;
 XX
 DR WPI; 1996-402134/40.
 DR N-PSDB; AAT36760.
 XX
 PT Direction of cellular immune response using therapeutic cell expressing 2
 PT chimeric receptors - comprising region binding to target cell and region
 PT that signals target cell destruction, or CD28 region, partic. for
 XX eliminating HIV-infected cells.
 PS Claim 7; Page 77-78; 120pp; English.
 XX
 CC A chimeric receptor (AAW0215) comprises the extracellular domain of an
 CC engineered form of the CD4 cellular receptor for HIV and the
 CC transmembrane and intracellular regions, including the cytoytic signal-
 CC transducing portion, of the mouse T-cell receptor eta chain. It can be
 CC obt'd. by inserting a gene fusion (AAT36760) into a vaccinia virus vector
 CC and expression in host cells. Chimeric receptors comprising CD4 fused to
 CC eta, eta (see also AAW02213) or FC receptor gamma (see also AAW02213)
 CC chains are capable of directing cytotoxic T lymphocytes to specifically
 CC recognise and kill cells expressing HIV gp120, thus providing a therapy
 CC for AIDS. (Updated on 16-OCT-2003 to standardise OS field)

XX Sequence 532 AA;
 SQ

QY Query Match 1.2%; Score 8; DB 2; Length 532;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
 |||||
 78 DRADSRRS 85

Db

RESULT 47
 AAW83141
 ID AAW83141 standard; protein; 532 AA.
 XX
 AC AAW83141;
 XX
 DT 03-FEB-1999 (first entry)
 DT
 XX
 DE Chimeric receptor containing human eta polypeptide.
 XX
 KW Human; zeta; eta; gamma; membrane-bound chimeric receptor; infection;
 KW tumour; cancer cell; autoimmunity-generated cell; T cell receptor; CD3;
 KW CD4; B cell receptor; FC receptor; pathogen; bacterial; fungal;
 KW protozoan; viral.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN US843728-A.
 PN
 PD 01-DEC-1998.
 PD
 PF 05-APR-1995; 95US-00417495.
 PF
 XX 07-MAR-1991; 91US-00655961.
 PR 06-MAR-1992; 92US-00847566.
 PR 28-FEB-1994; 94US-00203866.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Romeo C, Kolanus W, Seed B;
 XX
 DR WPI; 1999-044582/04.
 DR N-PSDB; AAV70157.
 XX

PT Membrane-bound chimeric receptors - comprising extracellular portion
 PT which recognises and binds a target cell and an intracellular portion of
 PT e.g. a T-cell receptor.
 XX
 PS Claim 11; Col 45-48; 57pp; English.
 CC The present invention describes DNA encoding a membrane-bound chimeric
 CC receptor comprising: (a) an extracellular portion that specifically
 CC recognises and binds a target cell or a target infective agent; and (b)
 CC an intracellular portion of a T-cell receptor CD3, zeta or eta
 CC polypeptide, a B-cell receptor polypeptide or an Fc receptor polypeptide.
 CC The present sequence represents a chimeric receptor containing the human
 CC eta polypeptide. Cells expressing chimeric receptors of the present
 CC invention can be administered to mammals in order to destroy pathogens
 CC (e.g. bacteria, fungi, protozoa or viruses, especially HIV), cancer cells
 CC or autoimmune-generated cells
 CC
 SQ Sequence 532 AA;
 QY Query Match 1.2%; Score 8; DB 2; Length 532;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 605 DRADSRSS 612
 78 DRADSRSS 85
 RESULT 48
 AAR26531
 ID AAR26531 standard; protein; 534 AA.
 AC AAR26531;
 XX
 XX 25-MAR-2003 (revised)
 DT 28-JAN-1993 (first entry)
 XX
 DE Sequence of CD4-IgG1 chimeric heavy chain heterotetramer.
 XX
 XX CD4-gamma 1 chimeric heavy chain homodimer; expression vector; HIV;
 KM therapy; diagnostic agent; inhibition.
 KW
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 205..302
 FT /label= CH1
 FT 303..317
 FT Region /label= hinge
 FT 318..427
 FT /label= CH2
 FT 428..534
 FT Region /label= CH3
 FT
 FT
 XX
 PN WO9213559-A1.
 XX
 PD 20-AUG-1992.
 XX
 PF 10-FEB-1992; 92WO-US001152.
 XX
 PR 08-FEB-1991; 91US-00654205.
 XX
 PA (PROG-) PROGENICS PHARM INC.
 XX
 PI Beaudry GA, Maddon PJ;
 XX
 DR WPI; 1992-299758/36.
 DR N-PSDB; AAQ27831.
 XX
 PT CD4-gamma 1 chimeric heavy chain homo-dimer and its expression vector -
 PT for preventing and treating HIV infection useful as a diagnostic agent.
 XX
 PS Example; Fig 4; 88pp; English.

XX
 CC The human CD4 cDNA is excised from the plasmid pSPeT4 and cloned into
 CC M13mp18. In order to excise a fragment containing the CH1 exon of the
 CC human gamma 1 heavy chain gene, the plasmid pBR gamma 1 is digested with
 CC SacII, and the SacII sites are then made flush using T4 DNA polymerase.
 CC The fragment containing the CH1 exon is then purified and ligated to the
 CC M13mp18(Cd4) vector. Oligonucleotide-mediated site-directed mutagenesis
 CC is then performed to juxtapose the CD4 and CH1 sequences in frame. The
 CC CD4-CH1 chimeric gene is then linearized and ligated to the pSL1-Pac1 DNA
 CC fragment of the plasmid pBR gamma 1 containing the hinge, CH2, and CH3
 CC exons of the human gamma 1 chain gene designated CD4-IgG1HC-pRCMV (ATCC
 CC 75192). (Updated on 25-MAR-2003 to correct PN field.)
 CC
 SQ Sequence 534 AA;
 QY Query Match 1.2%; Score 8; DB 2; Length 534;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 605 DRADSRSS 612
 78 DRADSRSS 85
 RESULT 49
 AAR04920
 ID AAR04920 standard; protein; 549 AA.
 AC AAR04920;
 XX
 XX 31-OCT-2002 (revised)
 DT 02-OCT-1990 (first entry)
 XX
 DE Immunoprotein PEX46.
 XX
 XX Soluble T4 protein; immunotoxin; Pseudomona endotoxin A; AIDS; HIV; ARC;
 KM angiosenin; fusion protein; PEX46; PEX46.
 KW
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 2..182
 FT /label= AAs 3-183 of T4 protein
 FT 183..549
 FT Region /label= AAs 253-613 of Pseudomonas exotoxin A (PE40)
 FT
 FT
 PN WO9004414-A.
 XX
 PD 03-MAY-1990.
 XX
 PF 18-OCT-1988; 88US-00259355.
 XX
 PR 18-OCT-1988; 88US-00259355.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 PI Meade HM, Lobb RR, Gates LL, Winkler G;
 XX
 DR WPI; 1990-163876/21.
 XX
 PT New immunotoxin contg. soluble T4 protein components and toxin - esp.
 PT Pseudomonas endotoxin A, for treating or controlling AIDS and related
 PT conditions, and new DNA sequences.
 XX
 PS Disclosure; Page ?; -pp; English.
 XX
 CC This fusion immunoprotein was produced by constructing a hybrid DNA
 CC sequence of a soluble T4 protein and a fragment of Pseudomonas exotoxin A
 CC (PE40) in which all binding region of PE was removed and which did not
 CC contain the boundary region. The hybrid DNA can then be inserted into an
 CC expression vector and used to produce recombinant fusion protein which is
 CC useful for preventing or treating AIDS, ARC, and HIV infections. The T4
 CC protein is an HIV receptor which binds to the virus or to infected cells

CC carrying the gp120/160 marker antigen, so provides v. specific targeting
 CC with minimal damage to non-target cells. Unmodified chain of PEX45 has a
 CC Mol. Wt. of 59,658. (Updated on 31-OCT-2002 to add missing OS field.)

XX SQ Sequence 549 AA;

Query Match 1.2%; Score 8; DB 2; Length 549;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
 |||||
 |||||

DB 54 DRADSRRS 61

RESULT 50

AAR04923 standard; protein; 557 AA.

XX AAR04923;

DT 31-OCT-2002 (revised)

DT 02-OCT-1990 (first entry)

DE Immunoprotein TANG11.

KM Soluble T4 protein; immunotoxin; Pseudomonas endotoxin A; AIDS; HIV; ARC;
 KM angiogenin; fusion protein; TANG11.

OS Synthetic.

XX FH Key Location/Qualifiers

FT Region 1..182
 /label= AAs 3-183 of T4 protein

FT Region 184..300
 /label= translocation domain of PE

FT Region 301..340
 /label= binding domain of PE

FT Region 341..429
 /label= linker from ADP ribosylation domain of PE

FT Region 431..559

FT Region /label= mature angiogenin

PN WO9004414-A.

PD 03-MAY-1990.

PF 18-OCT-1988; 88US-00259355.

PR 18-OCT-1988; 88US-00259355.

PA (BIOJ) BIOGEN INC.

PI Meade HM, Lobb RR, Gates LT, Winkler G;

XX WPI; 1990-163876/21.

PT New immunotoxin contg. soluble T4 protein components and toxin - esp.

PT Pseudomonas endotoxin A, for treating or controlling AIDS and related

PT conditions, and new DNA sequences.

XX PS Disclosure; Page ?; -pp; English.

CC This fusion immunoprotein was produced by constructing a hybrid DNA

CC sequence of a soluble T4 protein, a fragment of Pseudomonas exotoxin A

CC including the translocation domain and part of the ADP ribosylation

CC domain, and the gene for mature angiogenin. The hybrid DNA can then be

CC inserted into an expression vector and used to produce recombinant fusion

CC protein which is useful for preventing or treating AIDS, ARC, and HIV

CC infections. The T4 protein is an HIV receptor which binds to the virus or

CC to infected cells carrying the gp120/160 marker Ag, so provides v.

CC specific targeting with minimal damage to non-target cells. (Updated on

CC 31-OCT-2002 to add missing OS field.)

XX SQ Sequence 557 AA;

Query Match 1.2%; Score 8; DB 2; Length 557;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
 |||||
 |||||

DB 54 DRADSRRS 61

Search completed: June 21, 2004, 13:26:55
 Job time : 61 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 21, 2004, 13:26:59 ; Search time 47 Seconds
(without alignments)
3994.438 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665

Sequence: 1 MAHEMIGTQIVTERLVALL.....LGNVSSQSSFSGSMETIEVS 665

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1163542 seqs, 28213646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 174829

Minimum DB seq length: 473

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09A_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	71.0	517	9 US-09-964-277-21	Sequence 21, Appl1
2	472	71.0	665	9 US-09-816-494-2	Sequence 2, Appl1
3	472	71.0	665	9 US-09-964-277-2	Sequence 2, Appl1
4	472	71.0	665	12 US-10-072-012-680	Sequence 680, App
5	472	71.0	665	12 US-10-072-012-681	Sequence 681, App
6	472	71.0	665	12 US-10-168-506-14	Sequence 14, Appl1
7	472	71.0	665	12 US-10-343-357-7	Sequence 7, Appl1
8	472	71.0	665	15 US-10-094-749-2312	Sequence 2312, Ap
9	472	71.0	665	15 US-10-377-072-26	Sequence 26, Appl1
10	472	71.0	665	16 US-10-257-026-2	Sequence 2, Appl1
11	472	71.0	665	16 US-10-648-593-240	Sequence 240, App
12	472	71.0	665	16 US-10-648-593-247	Sequence 247, App
13	472	71.0	690	12 US-10-072-012-679	Sequence 679, App
14	472	71.0	690	12 US-10-072-012-703	Sequence 703, App
15	472	71.0	690	12 US-10-425-114-54204	Sequence 54204, A

16	394	59.2	680	12 US-10-072-012-256	Sequence 256, App
17	304	45.7	662	12 US-10-072-012-258	Sequence 258, App
18	258	38.8	672	12 US-10-296-115-1259	Sequence 1259, Ap
19	67	10.1	660	12 US-10-072-012-682	Sequence 682, App
20	64	9.6	677	12 US-10-072-012-683	Sequence 683, App
21	20	3.0	501	12 US-10-072-012-702	Sequence 702, App
22	20	3.0	625	12 US-10-072-012-699	Sequence 699, App
23	20	3.0	663	12 US-10-072-012-700	Sequence 700, App
24	12	1.8	616	12 US-10-072-012-266	Sequence 266, App
25	8	1.2	530	8 US-08-485-163-5	Sequence 5, Appl1
26	8	1.2	530	8 US-09-766-995-4	Sequence 4, Appl1
27	8	1.2	532	10 US-09-939-537-6	Sequence 6, Appl1
28	8	1.2	532	11 US-09-243-008-6	Sequence 6, Appl1
29	8	1.2	544	14 US-10-156-761-11387	Sequence 11387, A
30	8	1.2	575	10 US-09-939-537-4	Sequence 4, Appl1
31	8	1.2	575	11 US-09-243-008-4	Sequence 4, Appl1
32	8	1.2	590	9 US-09-934-060A-13	Sequence 13, Appl1
33	8	1.2	720	9 US-09-934-060A-2	Sequence 2, Appl1
34	8	1.2	720	9 US-09-934-060A-4	Sequence 4, Appl1
35	8	1.2	774	12 US-10-424-559-25796	Sequence 25796, A
36	8	1.2	788	14 US-10-073-118-26	Sequence 26, Appl1
37	7	1.1	475	12 US-10-282-122A-47101	Sequence 47101, A
38	7	1.1	478	12 US-10-424-559-245600	Sequence 245600, A
39	7	1.1	479	12 US-10-424-559-209874	Sequence 209874, A
40	7	1.1	498	14 US-10-032-585-7589	Sequence 7589, Ap
41	7	1.1	503	12 US-10-425-114-62654	Sequence 62654, A
42	7	1.1	512	16 US-10-355-238-13	Sequence 13, Appl1
43	7	1.1	512	16 US-10-355-238-23	Sequence 23, Appl1
44	7	1.1	515	16 US-10-355-238-14	Sequence 14, Appl1
45	7	1.1	515	16 US-10-355-238-25	Sequence 25, Appl1
46	7	1.1	516	14 US-10-156-761-8798	Sequence 8798, Ap
47	7	1.1	523	12 US-10-072-012-732	Sequence 732, App
48	7	1.1	525	12 US-10-282-122A-61404	Sequence 61404, A
49	7	1.1	526	14 US-10-361-460-73	Sequence 73, Appl1
50	7	1.1	532	14 US-10-289-757-106	Sequence 106, Appl1
51	7	1.1	537	12 US-10-424-559-164996	Sequence 164996, A
52	7	1.1	539	9 US-09-771-208-4	Sequence 4, Appl1
53	7	1.1	539	12 US-10-424-559-264414	Sequence 264414, A
54	7	1.1	542	14 US-10-289-757-107	Sequence 107, App
55	7	1.1	543	14 US-10-289-757-183	Sequence 183, App
56	7	1.1	544	12 US-10-282-122A-48564	Sequence 48564, A
57	7	1.1	547	12 US-10-425-114-50980	Sequence 50980, A
58	7	1.1	548	12 US-10-282-122A-65704	Sequence 65704, A
59	7	1.1	550	12 US-10-425-114-47928	Sequence 47928, A
60	7	1.1	555	9 US-09-764-868-878	Sequence 878, App
61	7	1.1	559	15 US-10-369-493-19468	Sequence 19468, A
62	7	1.1	560	10 US-09-863-776-69	Sequence 69, Appl1
63	7	1.1	561	9 US-09-764-864-1539	Sequence 1539, Ap
64	7	1.1	563	12 US-10-425-114-55259	Sequence 55259, A
65	7	1.1	564	9 US-09-999-248-14	Sequence 14, Appl1
66	7	1.1	564	12 US-09-878-722-245	Sequence 245, App
67	7	1.1	564	12 US-09-904-456-245	Sequence 245, App
68	7	1.1	564	14 US-10-157-031-176	Sequence 176, App
69	7	1.1	564	14 US-10-318-906A-2	Sequence 2, Appl1
70	7	1.1	564	14 US-10-319-236A-2	Sequence 2, Appl1
71	7	1.1	564	16 US-10-618-839-2	Sequence 2, Appl1
72	7	1.1	569	12 US-10-425-114-39496	Sequence 39496, A
73	7	1.1	580	14 US-10-156-761-12566	Sequence 12566, A
74	7	1.1	581	12 US-10-425-114-56980	Sequence 56980, A
75	7	1.1	584	14 US-10-029-386-32446	Sequence 32446, A

ALIGNMENTS

RESULT 1
US-09-964-277-21
; Sequence 21, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE

FILE REFERENCE: 200125.434
CURRENT APPLICATION NUMBER: US/09/964,277
CURRENT FILING DATE: 2001-09-25
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 21
LENGTH: 517
TYPE: PRT
ORGANISM: Homo sapiens
US-09-964-277-21

Query Match 71.0%; Score 472; DB 9; Length 517;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 194 TCPEPDPIPSHFLRVVNDSPCEKILPWLDKSVDFIEKAKASNGCVLVHCLAGISRAT 253
DB TCPEPDPIPSHFLRVVNDSPCEKILPWLDKSVDFIEKAKASNGCVLVHCLAGISRAT 105
QY 254 IAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKNOTGASGPKSKL 313
DB IAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKNOTGASGPKSKL 165
QY 314 LHLEKNEPVPAVSEGQKSETPLSPCADSATSEAAQRPVHPASVPSVQPSLLED 373
DB LHLEKNEPVPAVSEGQKSETPLSPCADSATSEAAQRPVHPASVPSVQPSLLED 225
QY 374 SPLVQALSGHLASADRLSDSNKLRSPSLDIKSVSYASMAASLHGSSSEDALEYKPS 433
DB SPLVQALSGHLASADRLSDSNKLRSPSLDIKSVSYASMAASLHGSSSEDALEYKPS 285
QY 434 TTLDTGNKLCQFSPVQELSEQTPETSPDKEASIPKKLQTPARPSDSQSKLHSVRTSSG 493
DB TTLDTGNKLCQFSPVQELSEQTPETSPDKEASIPKKLQTPARPSDSQSKLHSVRTSSG 345
QY 494 TAQRSLSPLRHSGVEDNYHTSFLGLSTSQOHLTKSAGLGKMHSDIILAPQSTPSL 553
DB TAQRSLSPLRHSGVEDNYHTSFLGLSTSQOHLTKSAGLGKMHSDIILAPQSTPSL 405
QY 554 TSSWYFATESHFFYSASAIYGSASYSAYSCQLPTCGDQVYSVRROKPSDRADRSRW 613
DB TSSWYFATESHFFYSASAIYGSASYSAYSCQLPTCGDQVYSVRROKPSDRADRSRW 465
QY 614 HESPEKQFKRRSCQMEFGESIMSENRSREELGKVGSSSFGSMELIEVS 665
DB HESPEKQFKRRSCQMEFGESIMSENRSREELGKVGSSSFGSMELIEVS 517
```

RESULT 2

US-09-816-494-2
Sequence 2, Application US/09816494
Patent No. US20020034807A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USBS THEREFOR
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,858
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-09-816-494-2

Query Match 71.0%; Score 472; DB 9; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 194 TCPEPDPIPSHFLRVVNDSPCEKILPWLDKSVDFIEKAKASNGCVLVHCLAGISRAT 253
DB TCPEPDPIPSHFLRVVNDSPCEKILPWLDKSVDFIEKAKASNGCVLVHCLAGISRAT 253
QY 254 IAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKNOTGASGPKSKL 313
DB IAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKNOTGASGPKSKL 313
QY 314 LHLEKNEPVPAVSEGQKSETPLSPCADSATSEAAQRPVHPASVPSVQPSLLED 373
DB LHLEKNEPVPAVSEGQKSETPLSPCADSATSEAAQRPVHPASVPSVQPSLLED 373
QY 374 SPLVQALSGHLASADRLSDSNKLRSPSLDIKSVSYASMAASLHGSSSEDALEYKPS 433
DB SPLVQALSGHLASADRLSDSNKLRSPSLDIKSVSYASMAASLHGSSSEDALEYKPS 433
QY 434 TTLDTGNKLCQFSPVQELSEQTPETSPDKEASIPKKLQTPARPSDSQSKLHSVRTSSG 493
DB TTLDTGNKLCQFSPVQELSEQTPETSPDKEASIPKKLQTPARPSDSQSKLHSVRTSSG 493
QY 494 TAQRSLSPLRHSGVEDNYHTSFLGLSTSQOHLTKSAGLGKMHSDIILAPQSTPSL 553
DB TAQRSLSPLRHSGVEDNYHTSFLGLSTSQOHLTKSAGLGKMHSDIILAPQSTPSL 553
QY 554 TSSWYFATESHFFYSASAIYGSASYSAYSCQLPTCGDQVYSVRROKPSDRADRSRW 613
DB TSSWYFATESHFFYSASAIYGSASYSAYSCQLPTCGDQVYSVRROKPSDRADRSRW 613
QY 614 HESPEKQFKRRSCQMEFGESIMSENRSREELGKVGSSSFGSMELIEVS 665
DB HESPEKQFKRRSCQMEFGESIMSENRSREELGKVGSSSFGSMELIEVS 665
```

RESULT 3

US-09-964-277-2
Sequence 2, Application US/09964277
Patent No. US2002013710A1
GENERAL INFORMATION:
APPLICANT: Luche, Ralf M.
TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.434
CURRENT APPLICATION NUMBER: US/09/964,277
CURRENT FILING DATE: 2001-09-25
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-09-964-277-2

Query Match 71.0%; Score 472; DB 9; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 194 TCPEPDPIPSHFLRVVNDSPCEKILPWLDKSVDFIEKAKASNGCVLVHCLAGISRAT 253
DB TCPEPDPIPSHFLRVVNDSPCEKILPWLDKSVDFIEKAKASNGCVLVHCLAGISRAT 253
QY 254 IAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKNOTGASGPKSKL 313
DB IAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKNOTGASGPKSKL 313
QY 314 LHLEKNEPVPAVSEGQKSETPLSPCADSATSEAAQRPVHPASVPSVQPSLLED 373
DB LHLEKNEPVPAVSEGQKSETPLSPCADSATSEAAQRPVHPASVPSVQPSLLED 373
QY 374 SPLVQALSGHLASADRLSDSNKLRSPSLDIKSVSYASMAASLHGSSSEDALEYKPS 433
DB SPLVQALSGHLASADRLSDSNKLRSPSLDIKSVSYASMAASLHGSSSEDALEYKPS 433
QY 434 TTLDTGNKLCQFSPVQELSEQTPETSPDKEASIPKKLQTPARPSDSQSKLHSVRTSSG 493
```

```
Db 434 TTLDGNTKLCQFSPVOELSEQTPETSPDKEBASIPKKLQOTARPDSQSKRLHSVRTSSSG 493
Qy 494 TAQRSLSLPLHRSGSVEDNYHTSFLGLSTSOOHLTKSAGLKGWHSIDLAPQISTPBL 553
Db 494 TAQRSLSLPLHRSGSVEDNYHTSFLGLSTSOOHLTKSAGLKGWHSIDLAPQISTPBL 553
Qy 554 TSSWYFATESSHFYSAIYGGASAYSAYSCQLPTCGQOYVSVRROKPSDRADRRSM 613
Db 554 TSSWYFATESSHFYSAIYGGASAYSAYSCQLPTCGQOYVSVRROKPSDRADRRSM 613
Qy 614 HESPEKQFKRRSCOMERGESIMSENREELGKVGSGSFGSGMEIIEVS 665
Db 614 HESPEKQFKRRSCOMERGESIMSENREELGKVGSGSFGSGMEIIEVS 665

RESULT 4
US-10-072-012-680
; Sequence 680, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zehnusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shinkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Baha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Raselli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 680
; LENGTH: 665
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-680
Query Match 71.0%; Score 472; DB 12; Length.665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 TCPKPDFIESHFLRVPVNDSCFKLPLWLDSDVPFIEKAKSNGCVLHYHCLAGISRST 253
Db 194 TCPKPDFIESHFLRVPVNDSCFKLPLWLDSDVPFIEKAKSNGCVLHYHCLAGISRST 253
Qy 254 IAIAYMKMDNSLDEAYRFEVKEKPTTSPNPNFGLQLDYEKIKNOTGASGPKSLKL 313
Db 254 IAIAYMKMDNSLDEAYRFEVKEKPTTSPNPNFGLQLDYEKIKNOTGASGPKSLKL 313
Qy 314 LHLKNEBPVPAVSEGGQKSETPLSPCADSATSEAGORPVHPASVPSPVQPSLBD 373
Db 314 LHLKNEBPVPAVSEGGQKSETPLSPCADSATSEAGORPVHPASVPSPVQPSLBD 373
Qy 374 SPLVQALSGIHSADRLSDSNLTKSFSLDITSVYSASMAASLHGFSSEDALEYKPS 433
Db 374 SPLVQALSGIHSADRLSDSNLTKSFSLDITSVYSASMAASLHGFSSEDALEYKPS 433
Qy 434 TTLDGNTKLCQFSPVOELSEQTPETSPDKEBASIPKKLQOTARPDSQSKRLHSVRTSSSG 493
Db 434 TTLDGNTKLCQFSPVOELSEQTPETSPDKEBASIPKKLQOTARPDSQSKRLHSVRTSSSG 493
Qy 494 TAQRSLSLPLHRSGSVEDNYHTSFLGLSTSOOHLTKSAGLKGWHSIDLAPQISTPBL 553
Db 494 TAQRSLSLPLHRSGSVEDNYHTSFLGLSTSOOHLTKSAGLKGWHSIDLAPQISTPBL 553
Qy 554 TSSWYFATESSHFYSAIYGGASAYSAYSCQLPTCGQOYVSVRROKPSDRADRRSM 613
Db 554 TSSWYFATESSHFYSAIYGGASAYSAYSCQLPTCGQOYVSVRROKPSDRADRRSM 613
Qy 614 HESPEKQFKRRSCOMERGESIMSENREELGKVGSGSFGSGMEIIEVS 665
Db 614 HESPEKQFKRRSCOMERGESIMSENREELGKVGSGSFGSGMEIIEVS 665

RESULT 5
US-10-072-012-681
; Sequence 681, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zehnusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shinkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Baha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Raselli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
```

```

; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 681
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-681

```

```

Query Match      71.0%; Score 472; DB 12; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 194 TCPKPDPIPESHFLRPVNDSPCEKILPMIDKSVDFIEKAKANGCVLHCLAGISRSAT 253
DB 194 TCPKPDPIPESHFLRPVNDSPCEKILPMIDKSVDFIEKAKANGCVLHCLAGISRSAT 253
QY 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKNOTGASGPKSKLKL 313
DB 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKNOTGASGPKSKLKL 313
QY 314 LHEKNEBPVAVSEGGQKSETPSPCADSATSEAGORPVHPASVPSVPSVPSLLED 373
DB 314 LHEKNEBPVAVSEGGQKSETPSPCADSATSEAGORPVHPASVPSVPSVPSLLED 373
QY 374 SPLVQALSGHLISADRLSDSNKLKRSFSLDIKSVSYASMAASLHGFSSEDALEYKPS 433
DB 374 SPLVQALSGHLISADRLSDSNKLKRSFSLDIKSVSYASMAASLHGFSSEDALEYKPS 433
QY 434 TLLDGTNKLCOFSPVOELSEQTPETSPDKEASIPKQLQTARPSDSOSKRLHSVRTSSSG 493
DB 434 TLLDGTNKLCOFSPVOELSEQTPETSPDKEASIPKQLQTARPSDSOSKRLHSVRTSSSG 493
QY 494 TAORSLSPLHRSGSVEDNYHTSFLFGLSTSOQHLTKSAGLGLKGMHSDIILAPQSTPSL 553
DB 494 TAORSLSPLHRSGSVEDNYHTSFLFGLSTSOQHLTKSAGLGLKGMHSDIILAPQSTPSL 553
QY 554 TSSWYFATESSHFYASAIYGGASAYASCQSLPTCGDQVYSVRRQKPSDRADSRSSW 613
DB 554 TSSWYFATESSHFYASAIYGGASAYASCQSLPTCGDQVYSVRRQKPSDRADSRSSW 613
QY 614 HESPEPEKQFKRRSCQMEFGESIMSENRREELGKVGSSQSFSGSMETIIEVS 665
DB 614 HESPEPEKQFKRRSCQMEFGESIMSENRREELGKVGSSQSFSGSMETIIEVS 665

```

```

RESULT 6
US-10-168-506-14
; Sequence 14, Application US/10168506
; Publication No. US20040053229A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: MARTINEZ, RICARDO

```

```

; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: HILL, RON
; APPLICANT: FLANAGAN, PETER
; TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
; FILE REFERENCE: 036602/1351
; CURRENT APPLICATION NUMBER: US/10/168,506
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: PCT/US00/34736
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-168-506-14

```

```

Query Match      71.0%; Score 472; DB 12; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 194 TCPKPDPIPESHFLRPVNDSPCEKILPMIDKSVDFIEKAKANGCVLHCLAGISRSAT 253
DB 194 TCPKPDPIPESHFLRPVNDSPCEKILPMIDKSVDFIEKAKANGCVLHCLAGISRSAT 253
QY 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKNOTGASGPKSKLKL 313
DB 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKIKNOTGASGPKSKLKL 313
QY 314 LHEKNEBPVAVSEGGQKSETPSPCADSATSEAGORPVHPASVPSVPSVPSLLED 373
DB 314 LHEKNEBPVAVSEGGQKSETPSPCADSATSEAGORPVHPASVPSVPSVPSLLED 373
QY 374 SPLVQALSGHLISADRLSDSNKLKRSFSLDIKSVSYASMAASLHGFSSEDALEYKPS 433
DB 374 SPLVQALSGHLISADRLSDSNKLKRSFSLDIKSVSYASMAASLHGFSSEDALEYKPS 433
QY 434 TLLDGTNKLCOFSPVOELSEQTPETSPDKEASIPKQLQTARPSDSOSKRLHSVRTSSSG 493
DB 434 TLLDGTNKLCOFSPVOELSEQTPETSPDKEASIPKQLQTARPSDSOSKRLHSVRTSSSG 493
QY 494 TAORSLSPLHRSGSVEDNYHTSFLFGLSTSOQHLTKSAGLGLKGMHSDIILAPQSTPSL 553
DB 494 TAORSLSPLHRSGSVEDNYHTSFLFGLSTSOQHLTKSAGLGLKGMHSDIILAPQSTPSL 553
QY 554 TSSWYFATESSHFYASAIYGGASAYASCQSLPTCGDQVYSVRRQKPSDRADSRSSW 613
DB 554 TSSWYFATESSHFYASAIYGGASAYASCQSLPTCGDQVYSVRRQKPSDRADSRSSW 613
QY 614 HESPEPEKQFKRRSCQMEFGESIMSENRREELGKVGSSQSFSGSMETIIEVS 665
DB 614 HESPEPEKQFKRRSCQMEFGESIMSENRREELGKVGSSQSFSGSMETIIEVS 665

```

```

RESULT 7
US-10-343-357-7
; Sequence 7, Application US/10343357
; Publication No. US20040058341A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y.Tom
; APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
; APPLICANT: LEE, Ernestine A.; HAPALIA, April U.A.
; APPLICANT: LU, Dying Aina M.; TRIBOUDEY, Catherine M.
; APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Henry; WARREN, Bridget A.
; APPLICANT: NGUYEN, Daniel B.; CHAWLA, Narinder K.
; APPLICANT: KEARNEY, Liam
; TITLE OF INVENTION: PROTEIN PHOSPHATASES

```

```
FILE REFERENCE: PI-0173 PCT
CURRENT APPLICATION NUMBER: US/10/343,357
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: PCT/US01/23716
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,679
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/223,272
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/224,309
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 60/226,728
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 60/229,254
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: US 60/231,366
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CDI
US-10-343-357-7
```

```
Query Match      71.0%; Score 472; DB 12; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 194 TCCKPDPFIESHFLRVPVNDSECEKILPWLKSVDFIEKAKANGCVLVHCLAGISRSAT 253
DB 194 TCCKPDPFIESHFLRVPVNDSECEKILPWLKSVDFIEKAKANGCVLVHCLAGISRSAT 253
QY 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKINQGTGASGPKSKLKL 313
DB 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKINQGTGASGPKSKLKL 313
QY 314 LHLKNEBPVPAVSEGGQSETPPLSPCCADATSEAAAGRPVHPASVPSPVQPSLLED 373
DB 314 LHLKNEBPVPAVSEGGQSETPPLSPCCADATSEAAAGRPVHPASVPSPVQPSLLED 373
QY 374 SPVQALSGHLHLSADRLSDSNKLRKSFSLDIKSVSYASMAASLHGFSSSEDALEYKPS 433
DB 374 SPVQALSGHLHLSADRLSDSNKLRKSFSLDIKSVSYASMAASLHGFSSSEDALEYKPS 433
QY 434 TTLDTGNKLCQFSPVQELSEQTPETSPPDKEASIPKKLQOTARPDSQSKRLHSVRTSSSG 493
DB 434 TTLDTGNKLCQFSPVQELSEQTPETSPPDKEASIPKKLQOTARPDSQSKRLHSVRTSSSG 493
QY 494 TAORSLSLPLHRSAGVEDNYHTSFLFGLSTSQOHLTKSAGLGKGMHSDILAPQSTPPL 553
DB 494 TAORSLSLPLHRSAGVEDNYHTSFLFGLSTSQOHLTKSAGLGKGMHSDILAPQSTPPL 553
QY 554 TSSWYPATESSHFYSAIYGGASAYSAYSCQLPTCCGQOVYVRRRQKPSDPAARRRW 613
DB 554 TSSWYPATESSHFYSAIYGGASAYSAYSCQLPTCCGQOVYVRRRQKPSDPAARRRW 613
QY 614 HESPEKQFKRRSCQMEFGESIMSENREBELGKVGSGSFFGSMETIEVS 665
DB 614 HESPEKQFKRRSCQMEFGESIMSENREBELGKVGSGSFFGSMETIEVS 665
```

```
RESULT 8
US-10-094-749-2312
; Sequence 2312, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
```

```
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOTYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASHIO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2312
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-10-094-749-2312
```

```
Query Match      71.0%; Score 472; DB 15; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 194 TCCKPDPFIESHFLRVPVNDSECEKILPWLKSVDFIEKAKANGCVLVHCLAGISRSAT 253
DB 194 TCCKPDPFIESHFLRVPVNDSECEKILPWLKSVDFIEKAKANGCVLVHCLAGISRSAT 253
QY 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKINQGTGASGPKSKLKL 313
DB 254 IAIAYIMKMDMSLDEAYRFVKEKRPPTISPNFNLGQLLDYEKKINQGTGASGPKSKLKL 313
QY 314 LHLKNEBPVPAVSEGGQSETPPLSPCCADATSEAAAGRPVHPASVPSPVQPSLLED 373
DB 314 LHLKNEBPVPAVSEGGQSETPPLSPCCADATSEAAAGRPVHPASVPSPVQPSLLED 373
QY 374 SPVQALSGHLHLSADRLSDSNKLRKSFSLDIKSVSYASMAASLHGFSSSEDALEYKPS 433
DB 374 SPVQALSGHLHLSADRLSDSNKLRKSFSLDIKSVSYASMAASLHGFSSSEDALEYKPS 433
QY 434 TTLDTGNKLCQFSPVQELSEQTPETSPPDKEASIPKKLQOTARPDSQSKRLHSVRTSSSG 493
DB 434 TTLDTGNKLCQFSPVQELSEQTPETSPPDKEASIPKKLQOTARPDSQSKRLHSVRTSSSG 493
QY 494 TAORSLSLPLHRSAGVEDNYHTSFLFGLSTSQOHLTKSAGLGKGMHSDILAPQSTPPL 553
DB 494 TAORSLSLPLHRSAGVEDNYHTSFLFGLSTSQOHLTKSAGLGKGMHSDILAPQSTPPL 553
QY 554 TSSWYPATESSHFYSAIYGGASAYSAYSCQLPTCCGQOVYVRRRQKPSDPAARRRW 613
DB 554 TSSWYPATESSHFYSAIYGGASAYSAYSCQLPTCCGQOVYVRRRQKPSDPAARRRW 613
QY 614 HESPEKQFKRRSCQMEFGESIMSENREBELGKVGSGSFFGSMETIEVS 665
DB 614 HESPEKQFKRRSCQMEFGESIMSENREBELGKVGSGSFFGSMETIEVS 665
```

```
RESULT 9
US-10-377-072-26
; Sequence 26, Application US/10377072
; Publication No. US20040009501A1
; GENERAL INFORMATION:
```

Db 554 TSSWFFATSSHHFYSASAIYGSASYSAYSCSLPTCDQYVYVRROKPSRADRSRW 613

QY 614 HEESEPFKQKRRSCOMFGEISIMSENRREELGVGSSQSFSGSMETIEVS 665

Db 614 HEESEPFKQKRRSCOMFGEISIMSENRREELGVGSSQSFSGSMETIEVS 665

RESULT 10
US-10-257-026-2

; Sequence 2, Application US/10257026
; Publication No. US2004008659A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: New dual specificity phosphatase
; FILE REFERENCE: DUSP10KWS
; CURRENT APPLICATION NUMBER: US/10/257,026
; CURRENT FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-257-026-2

Query Match 71.0%; Score 472; DB 16; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCPKPDPIPESHPLRVLPVNSFCCEKILPWLDKSVDFIEKAKANGCVLVHCLAGISRSAT 253

Db 194 TCPKPDPIPESHPLRVLPVNSFCCEKILPWLDKSVDFIEKAKANGCVLVHCLAGISRSAT 253

QY 254 IAIAYIMKRMDSGLDEAYRPFYKERRPTISPNFNLGQLDVEKKIKNOTGASGPKSKLKL 313

Db 254 IAIAYIMKRMDSGLDEAYRPFYKERRPTISPNFNLGQLDVEKKIKNOTGASGPKSKLKL 313

QY 314 LHEKPNBPVAVSEGGQKSETPSPPCADSATSEAGORPVHPASVPSPVPSLLED 373

Db 314 LHEKPNBPVAVSEGGQKSETPSPPCADSATSEAGORPVHPASVPSPVPSLLED 373

QY 374 SPLVQALSGHLISADRLSDSNKLRSPSLDIKSVSYASAMASLHGSSSEDALEYKPS 433

Db 374 SPLVQALSGHLISADRLSDSNKLRSPSLDIKSVSYASAMASLHGSSSEDALEYKPS 433

QY 434 TTLDTGNTKLCQFSYVQGLSECTPETSPPDKKEASIPKQLQTPARPSDSKRLHSVRTSSG 493

Db 434 TTLDTGNTKLCQFSYVQGLSECTPETSPPDKKEASIPKQLQTPARPSDSKRLHSVRTSSG 493

QY 494 TAQRSLSPILHRSGSVEDNYHTSFLFGLSTSOOHLTSAGLGLGWHSDILAPOTSTPSL 553

Db 494 TAQRSLSPILHRSGSVEDNYHTSFLFGLSTSOOHLTSAGLGLGWHSDILAPOTSTPSL 553

QY 554 TSSWFFATSSHHFYSASAIYGSASYSAYSCSLPTCDQYVYVRROKPSRADRSRW 613

Db 554 TSSWFFATSSHHFYSASAIYGSASYSAYSCSLPTCDQYVYVRROKPSRADRSRW 613

QY 614 HEESEPFKQKRRSCOMFGEISIMSENRREELGVGSSQSFSGSMETIEVS 665

Db 614 HEESEPFKQKRRSCOMFGEISIMSENRREELGVGSSQSFSGSMETIEVS 665

RESULT 11
US-10-648-593-240

; Sequence 240, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D02273 NP

CURRENT APPLICATION NUMBER: US/10/648,593
CURRENT FILING DATE: 2003-08-26
PRIOR APPLICATION NUMBER: 60/406,385
PRIOR FILING DATE: 2002-08-27
NUMBER OF SEQ ID NOS: 557
SOFTWARE: PatentIn version 3.2
SEQ ID NO 240
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-10-648-593-240

Query Match 71.0%; Score 472; DB 16; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCPCPDPIESHFLRVVNDVDFCEKILPMIDKSVDFIEKAKASNGCVLHVLGAGISRAT 253
DB 194 TCPCPDPIESHFLRVVNDVDFCEKILPMIDKSVDFIEKAKASNGCVLHVLGAGISRAT 253
QY 254 IAIAYIMKMDMSLDEAYRFVKEKRTTISPNNFQLQLDYEKKIKNOTGASGPKSKLXL 313
DB 254 IAIAYIMKMDMSLDEAYRFVKEKRTTISPNNFQLQLDYEKKIKNOTGASGPKSKLXL 313
QY 314 LHLEKNEPVPAVSEGGQKSETPPLSPCADSATSEAAAGRPVHPASVPSPVQPSLLBD 373
DB 314 LHLEKNEPVPAVSEGGQKSETPPLSPCADSATSEAAAGRPVHPASVPSPVQPSLLBD 373
QY 374 SPLVQALSGHLHSAADRLSDSNLTKRFSLDIKSVYSASMAASLHGFSSEDALEYYKRS 433
DB 374 SPLVQALSGHLHSAADRLSDSNLTKRFSLDIKSVYSASMAASLHGFSSEDALEYYKRS 433
QY 434 TLLDGTNKLCOFSPVQELSEQTPEPSPDKERASIPKKLQTPARPSDSQSKRLHSVRTSSSG 493
DB 434 TLLDGTNKLCOFSPVQELSEQTPEPSPDKERASIPKKLQTPARPSDSQSKRLHSVRTSSSG 493
QY 494 TAORSLSLPLHSGSVEDNYHTSFLFGLSTSOOHLTKSAGLKGWHSIDLAPQSTPSTL 553
DB 494 TAORSLSLPLHSGSVEDNYHTSFLFGLSTSOOHLTKSAGLKGWHSIDLAPQSTPSTL 553
QY 554 TSSWYPATESSHFYASAIYGSASYSAYSCSQLPTCGQVYSVRRROKPSDRADRRSM 613
DB 554 TSSWYPATESSHFYASAIYGSASYSAYSCSQLPTCGQVYSVRRROKPSDRADRRSM 613
QY 614 HESPEPEKQFKRRSCOMERGESIMSENREBELGKVGSSSFGSMETIEVS 665
DB 614 HESPEPEKQFKRRSCOMERGESIMSENREBELGKVGSSSFGSMETIEVS 665

RESULT 12

US-10-648-593-247
Sequence 247, Application US/10648593
Publication No. US20040106132A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
FILE REFERENCE: D0273 NP
CURRENT APPLICATION NUMBER: US/10/648,593
CURRENT FILING DATE: 2003-08-26
PRIOR APPLICATION NUMBER: 60/406,385
PRIOR FILING DATE: 2002-08-27
NUMBER OF SEQ ID NOS: 557
SOFTWARE: PatentIn version 3.2
SEQ ID NO 247
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-10-648-593-247

Query Match 71.0%; Score 472; DB 16; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCPCPDPIESHFLRVVNDVDFCEKILPMIDKSVDFIEKAKASNGCVLHVLGAGISRAT 253
DB 194 TCPCPDPIESHFLRVVNDVDFCEKILPMIDKSVDFIEKAKASNGCVLHVLGAGISRAT 253
QY 254 IAIAYIMKMDMSLDEAYRFVKEKRTTISPNNFQLQLDYEKKIKNOTGASGPKSKLXL 313
DB 254 IAIAYIMKMDMSLDEAYRFVKEKRTTISPNNFQLQLDYEKKIKNOTGASGPKSKLXL 313
QY 314 LHLEKNEPVPAVSEGGQKSETPPLSPCADSATSEAAAGRPVHPASVPSPVQPSLLBD 373
DB 314 LHLEKNEPVPAVSEGGQKSETPPLSPCADSATSEAAAGRPVHPASVPSPVQPSLLBD 373
QY 374 SPLVQALSGHLHSAADRLSDSNLTKRFSLDIKSVYSASMAASLHGFSSEDALEYYKRS 433
DB 374 SPLVQALSGHLHSAADRLSDSNLTKRFSLDIKSVYSASMAASLHGFSSEDALEYYKRS 433
QY 434 TLLDGTNKLCOFSPVQELSEQTPEPSPDKERASIPKKLQTPARPSDSQSKRLHSVRTSSSG 493
DB 434 TLLDGTNKLCOFSPVQELSEQTPEPSPDKERASIPKKLQTPARPSDSQSKRLHSVRTSSSG 493
QY 494 TAORSLSLPLHSGSVEDNYHTSFLFGLSTSOOHLTKSAGLKGWHSIDLAPQSTPSTL 553
DB 494 TAORSLSLPLHSGSVEDNYHTSFLFGLSTSOOHLTKSAGLKGWHSIDLAPQSTPSTL 553
QY 554 TSSWYPATESSHFYASAIYGSASYSAYSCSQLPTCGQVYSVRRROKPSDRADRRSM 613
DB 554 TSSWYPATESSHFYASAIYGSASYSAYSCSQLPTCGQVYSVRRROKPSDRADRRSM 613
QY 614 HESPEPEKQFKRRSCOMERGESIMSENREBELGKVGSSSFGSMETIEVS 665
DB 614 HESPEPEKQFKRRSCOMERGESIMSENREBELGKVGSSSFGSMETIEVS 665

RESULT 13

US-10-072-012-679
Sequence 679, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernyev, Valizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernusen, Bryan
APPLICANT: Paturajan, Meera
APPLICANT: Shimkete, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Esha
APPLICANT: Padigaru, Muraidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Molenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsbrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31

```

; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 679
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-679

```

```

Query Match      71.0%; Score 472; DB 12; Length 690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 194 TCPRPDIPIESHFLRVVNDSPFCEKILPWLKSVDFIEKAKASNGCVLVHCLAGISRSAT 253
DB 219 TCPRPDIPIESHFLRVVNDSPFCEKILPWLKSVDFIEKAKASNGCVLVHCLAGISRSAT 278
QY 254 IAIAYIMKMDMSLDIAVRYVKEKRPITISPNFNLGQLLDYEKKIKNOTGASGPKSKL 313
DB 279 IAIAYIMKMDMSLDIAVRYVKEKRPITISPNFNLGQLLDYEKKIKNOTGASGPKSKL 338
QY 314 LHLEKNEVPYPAVSEGGQKSETPPLSPPCADSATSEAAQGPVHPASVPSVQPSLLED 373
DB 339 LHLEKNEVPYPAVSEGGQKSETPPLSPPCADSATSEAAQGPVHPASVPSVQPSLLED 398
QY 374 SPLVQALSGHLISADRLSDNSNKLKRSFSLDIKSVSYASMAASLHGFSSEDALEYYKPS 433
DB 399 SPLVQALSGHLISADRLSDNSNKLKRSFSLDIKSVSYASMAASLHGFSSEDALEYYKPS 458
QY 434 TTLDTGNKLCQFSPVQELSEQTPETSPDKEEASIPKKLQIARPSDSQSKLHVSRTSSSG 493
DB 459 TTLDTGNKLCQFSPVQELSEQTPETSPDKEEASIPKKLQIARPSDSQSKLHVSRTSSSG 518
QY 494 TKQRSILSLPHRSGSVEDNYHTSFLFGLSTSQOHLTKSAGLGKGMHSDIAPQTSPTSL 553
DB 519 TKQRSILSLPHRSGSVEDNYHTSFLFGLSTSQOHLTKSAGLGKGMHSDIAPQTSPTSL 578
QY 554 TSSWYFATESHPSHYASAIYGGASASAYSCSOLPTCCGQVYSVRRQRPSPRARSRSRW 613
DB 579 TSSWYFATESHPSHYASAIYGGASASAYSCSOLPTCCGQVYSVRRQRPSPRARSRSRW 638
QY 614 HESPEPEKQFKRRSCOMEPGESIMSENRREELGKYGQSQSPSGMEITIEVS 665
DB 639 HESPEPEKQFKRRSCOMEPGESIMSENRREELGKYGQSQSPSGMEITIEVS 690

```

```

RESULT 14
US-10-072-012-703
; Sequence 703, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patuturajan, Meera
; APPLICANT: Shinkens, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara

```

```

; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taudier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 703
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-703

```

```

Query Match      71.0%; Score 472; DB 12; Length 690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 194 TCPRPDIPIESHFLRVVNDSPFCEKILPWLKSVDFIEKAKASNGCVLVHCLAGISRSAT 253
DB 219 TCPRPDIPIESHFLRVVNDSPFCEKILPWLKSVDFIEKAKASNGCVLVHCLAGISRSAT 278
QY 254 IAIAYIMKMDMSLDIAVRYVKEKRPITISPNFNLGQLLDYEKKIKNOTGASGPKSKL 313
DB 279 IAIAYIMKMDMSLDIAVRYVKEKRPITISPNFNLGQLLDYEKKIKNOTGASGPKSKL 338
QY 314 LHLEKNEVPYPAVSEGGQKSETPPLSPPCADSATSEAAQGPVHPASVPSVQPSLLED 373
DB 339 LHLEKNEVPYPAVSEGGQKSETPPLSPPCADSATSEAAQGPVHPASVPSVQPSLLED 398
QY 374 SPLVQALSGHLISADRLSDNSNKLKRSFSLDIKSVSYASMAASLHGFSSEDALEYYKPS 433
DB 399 SPLVQALSGHLISADRLSDNSNKLKRSFSLDIKSVSYASMAASLHGFSSEDALEYYKPS 458
QY 434 TTLDTGNKLCQFSPVQELSEQTPETSPDKEEASIPKKLQIARPSDSQSKLHVSRTSSSG 493
DB 459 TTLDTGNKLCQFSPVQELSEQTPETSPDKEEASIPKKLQIARPSDSQSKLHVSRTSSSG 518
QY 494 TKQRSILSLPHRSGSVEDNYHTSFLFGLSTSQOHLTKSAGLGKGMHSDIAPQTSPTSL 553

```


Db 519 TAQRSLSPHRSQVEDNHTSFLGSLSTSQHLTKSAGLKGWMSDILAQOTSTPSL 578
QY 554 TSSWYATSSSHYYSASAIYGSASYSAYSCSOLPTCGDQVYVRRRQKSPDADSRSSW 613
Db 579 TSSWYATSSSHYYSASAIYGSASYSAYSCSOLPTCGDQVYVRRRQKSPDADSRSSW 638
QY 614 HESSPEKOPKRRSCOMEFGESIMSENRSREELGKVGSSGSSFGSMELLIEVS 665
Db 639 HESSPEKOPKRRSCOMEFGESIMSENRSREELGKVGSSGSSFGSMELLIEVS 690

RESULT 15

US-10-425-114-54204
Sequence 54204, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovacic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21 (5311) B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 54204
LENGTH: 690
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Clone ID: LIB4119-028-H6_FIL.pep
US-10-425-114-54204

Query Match 71.0%; Score 472; DB 12; Length 690;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCPRPPIPSHFLRVVNDSPCEKILPMUDKSVDFIEKAKASNGCVLHLAGISRAT 253
Db 219 TCPRPPIPSHFLRVVNDSPCEKILPMUDKSVDFIEKAKASNGCVLHLAGISRAT 278
QY 254 IAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKNOTGASGPKSKL 313
Db 279 IAIAYIMKRMDSLDEAYRFVKEKRPITISPNFNLGQLDYEKKIKNOTGASGPKSKL 338
QY 314 LHEKNEPVPVASEGGQKSETPLSPCCADTSATSEAAQGPVHPASVPSVQPSLLE 373
Db 339 LHEKNEPVPVASEGGQKSETPLSPCCADTSATSEAAQGPVHPASVPSVQPSLLE 398
QY 374 SPLVQALSGHLGADRLSDNKLKRSFLDIKSVYSASMAAHLGFSSEDLLEYKPS 433
Db 399 SPLVQALSGHLGADRLSDNKLKRSFLDIKSVYSASMAAHLGFSSEDLLEYKPS 458
QY 434 TTLDTGNKLCOFSPVQELSFQTPETSPDKKEASIPKKLQTPARPSDSQSKLHSVRTSSG 493
Db 459 TTLDTGNKLCOFSPVQELSFQTPETSPDKKEASIPKKLQTPARPSDSQSKLHSVRTSSG 518
QY 494 TAQRSLSPHRSQVEDNHTSFLGSLSTSQHLTKSAGLKGWMSDILAQOTSTPSL 553
Db 519 TAQRSLSPHRSQVEDNHTSFLGSLSTSQHLTKSAGLKGWMSDILAQOTSTPSL 578
QY 554 TSSWYATSSSHYYSASAIYGSASYSAYSCSOLPTCGDQVYVRRRQKSPDADSRSSW 613
Db 579 TSSWYATSSSHYYSASAIYGSASYSAYSCSOLPTCGDQVYVRRRQKSPDADSRSSW 638
QY 614 HESSPEKOPKRRSCOMEFGESIMSENRSREELGKVGSSGSSFGSMELLIEVS 665
Db 639 HESSPEKOPKRRSCOMEFGESIMSENRSREELGKVGSSGSSFGSMELLIEVS 690

RESULT 16

US-10-072-012-256
Sequence 256, Application US/10072012
Publication No. US2004003493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernhusen, Bryan
APPLICANT: Paturajan, Weera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Beba
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimир Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsbrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-08
PRIOR APPLICATION NUMBER: 60/267,459
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 256
LENGTH: 680
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-256

Query Match 59.2%; Score 394; DB 12; Length 680;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 RFYKERPTISPNFNLGQLDYEKKIKNOTGASGPKSKLHLHEKNEPVPVASEGGQ 331
Db 287 RFYKERPTISPNFNLGQLDYEKKIKNOTGASGPKSKLHLHEKNEPVPVASEGGQ 346
QY 332 KSETPLSPCCADTSATSEAAQGPVHPASVPSVQPSLLEDSPLVQALSGHLGADRL 391
Db 347 KSETPLSPCCADTSATSEAAQGPVHPASVPSVQPSLLEDSPLVQALSGHLGADRL 406

QY 392 DSNKLRKSFSLDIKSVSYSSAMASLHGFSSSEDALEYKSTTLDTGNKLCQFSPVOEL 451
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patentin Ver. 2.1
Db 407 DSNKLRKSFSLDIKSVSYSSAMASLHGFSSSEDALEYKSTTLDTGNKLCQFSPVOEL 466
; SEQ ID NO 258
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Homo sapiens
QY 452 SEQTPEPSPKREASIPKKLQTARPSDSQSKRLHSVTSSTSSGTAQSRSLSPLRSGSVED 511
Db 467 SEQTPEPSPKREASIPKKLQTARPSDSQSKRLHSVTSSTSSGTAQSRSLSPLRSGSVED 526
QY 512 NYHTSFLFGLSTSQOHLTKSAGLKGWHSDDLAPQSTPSTLSSWYFATESHFFYSASA 571
Db 527 NYHTSFLFGLSTSQOHLTKSAGLKGWHSDDLAPQSTPSTLSSWYFATESHFFYSASA 586
QY 572 IYGSASYSASVSCSGLPTCCDQVYVRRQKPSDRADSRSMHEESPFEKQFKRRSCOME 631
Db 587 IYGSASYSASVSCSGLPTCCDQVYVRRQKPSDRADSRSMHEESPFEKQFKRRSCOME 646
QY 632 FGESIMSENRSREELGKVGSSQSFSGSMETIEVS 665
Db 647 FGESIMSENRSREELGKVGSSQSFSGSMETIEVS 680

RESULT 17

US-10-072-012-258
; Sequence 258, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patunrajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Beha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459

; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 258
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-258

Query Match 45.7%; Score 304; DB 12; Length 662;
Best Local Similarity 100.0%; Pred. No. 1,1e-280;
Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 SVPSVQPSLLEPSPPLVQALSGHLSDADLEDSNKLKRSFSLDIKSVTSASMAASLHGF 421
Db 359 SVPSVQPSLLEPSPPLVQALSGHLSDADLEDSNKLKRSFSLDIKSVTSASMAASLHGF 418
QY 422 SSEDALLEYKSTTLDTGNKLCQFSPVOELSBQTPETSPDKREASIPKKLQTARPSDSQS 481
Db 419 SSEDALLEYKSTTLDTGNKLCQFSPVOELSBQTPETSPDKREASIPKKLQTARPSDSQS 478
QY 482 KRLHSVTSSTSGTAQSRSLSPLRSGSVEDNYHTSFLFGLSTSQOHLTKSAGLKGWHS 541
Db 479 KRLHSVTSSTSGTAQSRSLSPLRSGSVEDNYHTSFLFGLSTSQOHLTKSAGLKGWHS 538
QY 542 DILAPQSTPSTLSSWYFATESHFFYSASATYGSASVSCSGLPTCCDQVYVRRQ 601
Db 539 DILAPQSTPSTLSSWYFATESHFFYSASATYGSASVSCSGLPTCCDQVYVRRQ 598
QY 602 KPSDRADSRSMHEESPFEKQFKRRSCOMEFGESIMSENRSREELGKVGSSQSFSGSMET 661
Db 599 KPSDRADSRSMHEESPFEKQFKRRSCOMEFGESIMSENRSREELGKVGSSQSFSGSMET 658
QY 662 IEVS 665
Db 659 IEVS 662

RESULT 18

US-10-296-115-1259
; Sequence 1259, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1259
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATRE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(672)
; OTHER INFORMATION: Xaa = any amino acid or other as shown in Table 3
US-10-296-115-1259

Query Match 38.8%; Score 258; DB 12; Length 672;
Best Local Similarity 100.0%; Pred. No. 7.8e-237;
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCPKPDFIPSHFLRVVUNSPCEKILPWLDKSVDFLEKAKASGCVLHCLIGISSAT 253
Db 201 TCPKPDFIPSHFLRVVUNSPCEKILPWLDKSVDFLEKAKASGCVLHCLIGISSAT 260
QY 254 IAIYIMKRMWMSLDEAYRVRKEKRPITSPNFVLGLLDYKKIKQTGASGPKSLKL 313

```

; TYPE: PRT
; ORGANISM: Mus musculus
US-10-072-012-682
Query Match      10.1%; Score 67; DB 12; Length 660;
Best Local Similarity 100.0%; Pred. No. 1.1e-54;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LVALLSGTEKYLIDSRPEVYNTSHILEAININSGKLMKRLQDDKYLITELIOHSK 74
DB 15 LVALLSGTEKYLIDSRPEVYNTSHILEAININSGKLMKRLQDDKYLITELIOHSK 74

QY 434 TTLDTGNKLCQFSPVOEL 451
DB 441 TTLDTGNKLCQFSPVOEL 458

RESULT 19
US-10-072-012-682
; Sequence 682, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shimkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muraidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie E.
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 682
; LENGTH: 660
```

```

; TYPE: PRT
; ORGANISM: Mus musculus
US-10-072-012-683
Query Match      10.1%; Score 67; DB 12; Length 660;
Best Local Similarity 100.0%; Pred. No. 1.1e-54;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LVALLSGTEKYLIDSRPEVYNTSHILEAININSGKLMKRLQDDKYLITELIOHSK 74
DB 15 LVALLSGTEKYLIDSRPEVYNTSHILEAININSGKLMKRLQDDKYLITELIOHSK 74

QY 75 HKVDIDC 81
DB 75 HKVDIDC 81

RESULT 20
US-10-072-012-683
; Sequence 683, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shimkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muraidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie E.
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 683
```

/ LENGTH: 677
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-10-072-012-683

Query Match 9.6%; Score 64; DB 12; Length 677;
Best Local Similarity 100.0%; Pred. No. 7,8e-52;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 SFSVHLGGFAPFCPCGKSTVPTCTISQCPVANIIGTRILPNLYGCGORD 173
DB 114 SFSVHLGGFAPFCPCGKSTVPTCTISQCPVANIIGTRILPNLYGCGORD 173

QY 174 VLNK 177
DB 174 VLNK 177

RESULT 21

US-10-072-012-702
/ Sequence 702, Application US/10072012
/ Publication No. US2004003493A1
/ GENERAL INFORMATION:

/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Paturajan, Meera
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangolli, Esna
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie E.
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Coleman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grose, William M.
/ APPLICANT: Alsobrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072,012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 60/266,767
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: 60/267,057
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/266,975
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/267,459
/ PRIOR FILING DATE: 2001-02-08
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1391
/ SOFTWARE: PatentIn Ver. 2.1

/ SEQ ID NO 702
/ LENGTH: 501
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: (469)
/ OTHER INFORMATION: wherein Xaa is any amino acid.
US-10-072-012-702

Query Match 3.0%; Score 20; DB 12; Length 501;
Best Local Similarity 100.0%; Pred. No. 5,4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 VHCLAGISRSATIAIAYIMK 261
DB 98 VHCLAGISRSATIAIAYIMK 117

RESULT 22

US-10-072-012-699
/ Sequence 699, Application US/10072012
/ Publication No. US2004003493A1
/ GENERAL INFORMATION:

/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Paturajan, Meera
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangolli, Esna
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie E.
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Coleman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grose, William M.
/ APPLICANT: Alsobrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072,012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 60/266,767
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: 60/267,057
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/266,975
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/267,459
/ PRIOR FILING DATE: 2001-02-08
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1391

```

; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 699
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-699

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 625;
Pred. No. 6,6e-10; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

QY      242 VHCLAGISRSATTAIAYIMK 261
      |||||
Db      244 VHCLAGISRSATTAIAYIMK 263

RESULT 23
US-10-072-012-700
; Sequence 700, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zehrusen, Bryan
; APPLICANT: Patuturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 700
; LENGTH: 663
```

```

; TYPE: PRT
; ORGANISM: Mus musculus
US-10-072-012-700

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 663;
Pred. No. 7e-10; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

QY      242 VHCLAGISRSATTAIAYIMK 261
      |||||
Db      244 VHCLAGISRSATTAIAYIMK 263

RESULT 24
US-10-072-012-266
; Sequence 266, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zehrusen, Bryan
; APPLICANT: Patuturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 266
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-266
```

Query Match 1.8%; Score 12; DB 12; Length 616;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 ISPNFNLGQL 292
|||||
DB 285 ISPNFNLGQL 296

RESULT 25
US-08-485-163-5
; Sequence 5, Application US/08485163
; Publication No. US2002098191A1
; GENERAL INFORMATION:
; APPLICANT: Beaudry, Gary A.
; TITLE OF INVENTION: CD4-GAMMA2 CD4-IGG2 CHIMERAS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,163
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 37690-II-1-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: homo sapien
; CELL TYPE: lymphocyte
US-08-485-163-5

Query Match 1.2%; Score 8; DB 8; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
|||||
DB 78 DRADSRRS 85

RESULT 26
US-09-766-995-4
; Sequence 4, Application US/09766995
; Patent No. US20020052481A1
; GENERAL INFORMATION:
; APPLICANT: Graham P. Allaway et al.
; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED CD4-GAMMA2 AND CD4-IGG2 IMMUNOCONJ
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 2048/41215-CB/JPW/SHS

CURRENT APPLICATION NUMBER: US/09/766,995
; CURRENT FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4

LENGTH: 530
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-766-995-4

Query Match 1.2%; Score 8; DB 9; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
|||||
DB 78 DRADSRRS 85

RESULT 27
US-09-939-537-6
; Sequence 6, Application US/09939537
; Publication No. US20030138410A1
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; Banapour, Babak
; Romeo, Charles
; Kolarus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Ebling LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,537
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Ebling, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-939-537-6

REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001

TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-939-537-4

Query Match 1.2%; Score 8; DB 10; Length 575;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
DB 78 DRADSRSS 85

RESULT 31
US-09-243-008-4
Sequence 4, Application US/09243008
Publication No. US2004000534A1
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor Chimeras
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/243,008
FILING DATE: 02-Feb-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/394,176
FILING DATE: SEPTEMBER 11, 1995
APPLICATION NUMBER: 08/203,866
FILING DATE: February 28, 1994
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Karen F. Lech, Ph.D.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/270001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-243-008-4

Query Match 1.2%; Score 8; DB 11; Length 575;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
DB 78 DRADSRSS 85

RESULT 32
US-09-934-060A-13
Sequence 13, Application US/09934060A
Patent No. US20020155121A1
GENERAL INFORMATION:
APPLICANT: Devico, Anthony L.
APPLICANT: Fouts, Timothy R.
APPLICANT: Tuskan, Robert G.
TITLE OF INVENTION: VIRUS COAT PROTEIN/RECEPTOR CHIMERAS AND METHODS OF USE
FILE REFERENCE: 4115-144 CIP
CURRENT APPLICATION NUMBER: US/09/934,060A
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 09/684,026
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/158,321
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 590
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthesized construct
NAME/KEY: MISC FEATURE
LOCATION: (586)..(586)
OTHER INFORMATION: Xaa can be any amino acid
NAME/KEY: MISC FEATURE
LOCATION: (589)..(589)
OTHER INFORMATION: Xaa can be any amino acid
US-09-934-060A-13

Query Match 1.2%; Score 8; DB 9; Length 590;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
DB 450 DRADSRSS 457

RESULT 33
US-09-934-060A-2
Sequence 2, Application US/09934060A
Patent No. US20020155121A1
GENERAL INFORMATION:
APPLICANT: Devico, Anthony L.
APPLICANT: Fouts, Timothy R.
APPLICANT: Tuskan, Robert G.
TITLE OF INVENTION: VIRUS COAT PROTEIN/RECEPTOR CHIMERAS AND METHODS OF USE
FILE REFERENCE: 4115-144 CIP
CURRENT APPLICATION NUMBER: US/09/934,060A
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 09/684,026
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/158,321
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 720
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Synthesized construct
NAME/KEY: MISC FEATURE
LOCATION: (716)..(716)
OTHER INFORMATION: Xaa can be any amino acid
NAME/KEY: MISC FEATURE
LOCATION: (719)..(719)
OTHER INFORMATION: Xaa can be any amino acid
US-09-934-060A-2

Query Match 1.2%; Score 8; DB 9; Length 720;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
DB 580 DRADSRRS 587

RESULT 34
US-09-934-060A-4
Sequence 4, Application US/09934060A
Patent No. US20020155121A1
GENERAL INFORMATION:
APPLICANT: Devico, Anthony L.
APPLICANT: Fouts, Timothy R.
APPLICANT: Tusken, Robert G.
TITLE OF INVENTION: VIRUS COAT PROTEIN/RECEPTOR CHIMERAS AND METHODS OF USE
FILE REFERENCE: 4115-144 CIP
CURRENT APPLICATION NUMBER: US/09/934,060A
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 09/684,026
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/158,321
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 720
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthesized construct
NAME/KEY: MISC FEATURE
LOCATION: (716)..(716)
OTHER INFORMATION: Xaa can be any amino acid
NAME/KEY: MISC FEATURE
LOCATION: (719)..(719)
OTHER INFORMATION: Xaa can be any amino acid
US-09-934-060A-4

Query Match 1.2%; Score 8; DB 9; Length 720;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
DB 580 DRADSRRS 587

RESULT 35
US-10-424-599-257796
Sequence 257796, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223) B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 257796
LENGTH: 774
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(774)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_74813C.1.pep
US-10-424-599-257796

Query Match 1.2%; Score 8; DB 12; Length 774;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 647 GRYGSSSS 654
DB 736 GRYGSSSS 743

RESULT 36
US-10-073-118-26.
Sequence 26, Application US/10073118
Publication No. US20030054554A1
GENERAL INFORMATION:
APPLICANT: BECQUART, JEROME
TITLE OF INVENTION: ALBUMIN DERIVATIVES WITH THERAPEUTIC FUNCTIONS
FILE REFERENCE: 06832.1429-03
CURRENT APPLICATION NUMBER: US/10/073,118
CURRENT FILING DATE: 2002-02-12
PRIOR APPLICATION NUMBER: 09/551,635
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 09/004,319
PRIOR FILING DATE: 1998-01-08
PRIOR APPLICATION NUMBER: 08/479,146
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 08/295,078
PRIOR FILING DATE: 1994-08-26
PRIOR APPLICATION NUMBER: 08/121,236
PRIOR FILING DATE: 1993-09-13
PRIOR APPLICATION NUMBER: 07/955,243
PRIOR FILING DATE: 1992-10-01
PRIOR APPLICATION NUMBER: 07/561,879
PRIOR FILING DATE: 1990-08-02
PRIOR APPLICATION NUMBER: FR 89 10480
PRIOR FILING DATE: 1989-08-03
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 26
LENGTH: 788
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Protein fusion
OTHER INFORMATION: prepro-HSA-VIIV2
US-10-073-118-26

Query Match 1.2%; Score 8; DB 14; Length 788;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
DB 662 DRADSRRS 669

RESULT 37
US-10-282-122A-47101
Sequence 47101, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:

```

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EPIRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47101
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-282-122A-47101

Query Match          1.1%; Score 7; DB 12; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      296 KKIKNOT 302
DB      469 KKIKNOT 475

RESULT 38
US-10-424-599-245600
; Sequence 245600, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 245600
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Glycine max
```

```

; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(478)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63809C.1.pcp
US-10-424-599-245600

Query Match          1.1%; Score 7; DB 12; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      518 LFGSLTS 524
DB      125 LFGSLTS 131

RESULT 39
US-10-424-599-209874
; Sequence 209874, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 209874
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(479)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_31545C.1.pcp
US-10-424-599-209874

Query Match          1.1%; Score 7; DB 12; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      518 LFGSLTS 524
DB      126 LFGSLTS 132

RESULT 40
US-10-032-585-7589
; Sequence 7589, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7589
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7589
```

Query Match 1.1%; Score 7; DB 14; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 489 TSSSGTA 495
|||||
DB 200 TSSSGTA 206

RESULT 41
US-10-425-114-62654
; Sequence 62654, Application US/10425114
; Publication No. US2004004888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62654
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3066-034-F12_F11.pep
US-10-425-114-62654

Query Match 1.1%; Score 7; DB 12; Length 503;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 FIEKAKA 235
|||||
DB 105 FIEKAKA 111

RESULT 42
US-10-355-238-13
; Sequence 13, Application US/10355238
; Publication No. US20040091987A1
; GENERAL INFORMATION:
; APPLICANT: SPELLIG, TITMAN
; TITLE OF INVENTION: PROCESS FOR THE OVEREXPRESSION OF DEHYDROGENASES
; FILE REFERENCE: SCH-1898
; CURRENT APPLICATION NUMBER: US/10/355,238
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: DE 10204798.1
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/382,569
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Brevibacterium maris
US-10-355-238-13

Query Match 1.1%; Score 7; DB 16; Length 512;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 EDALLEY 430
|||||
DB 72 EDALLEY 78

RESULT 43
US-10-355-238-23
; Sequence 23, Application US/10355238
; Publication No. US20040091987A1
; GENERAL INFORMATION:
; APPLICANT: SPELLIG, TITMAN
; TITLE OF INVENTION: PROCESS FOR THE OVEREXPRESSION OF DEHYDROGENASES
; FILE REFERENCE: SCH-1898
; CURRENT APPLICATION NUMBER: US/10/355,238
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: DE 10204798.1
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/382,569
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Brevibacterium maris
US-10-355-238-23

Query Match 1.1%; Score 7; DB 16; Length 512;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 EDALLEY 430
|||||
DB 72 EDALLEY 78

RESULT 44
US-10-355-238-14
; Sequence 14, Application US/10355238
; Publication No. US20040091987A1
; GENERAL INFORMATION:
; APPLICANT: SPELLIG, TITMAN
; TITLE OF INVENTION: PROCESS FOR THE OVEREXPRESSION OF DEHYDROGENASES
; FILE REFERENCE: SCH-1898
; CURRENT APPLICATION NUMBER: US/10/355,238
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: DE 10204798.1
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/382,569
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Arthrobacter simplex
US-10-355-238-14

Query Match 1.1%; Score 7; DB 16; Length 515;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 EDALLEY 430
|||||
DB 71 EDALLEY 77

RESULT 45
US-10-355-238-25
; Sequence 25, Application US/10355238
; Publication No. US20040091987A1
; GENERAL INFORMATION:
; APPLICANT: SPELLIG, TITMAN
; TITLE OF INVENTION: PROCESS FOR THE OVEREXPRESSION OF DEHYDROGENASES
; FILE REFERENCE: SCH-1898
; CURRENT APPLICATION NUMBER: US/10/355,238

```
/ CURRENT FILING DATE: 2003-01-31
/ PRIOR APPLICATION NUMBER: DE 10204798.1
/ PRIOR FILING DATE: 2002-02-01
/ PRIOR APPLICATION NUMBER: 60/382,569
/ PRIOR FILING DATE: 2002-05-24
/ NUMBER OF SEQ ID NOS: 29
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 25
/ LENGTH: 515
/ TYPE: PRT
/ ORGANISM: Arthrobacter simplex
US-10-355-238-25

Query Match
Best Local Similarity 100.0%; Score 7; DB 16; Length 515;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 EDALLEY 430
DB 71 EDALLEY 77

RESULT 46
US-10-156-761-8798
/ Sequence 8798, Application US/10156761
/ Publication No. US20030119018A1
/ GENERAL INFORMATION:
/ APPLICANT: OMTA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIBA, TADAYOSHI
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10/156,761
/ CURRENT FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 8798
/ LENGTH: 516
/ TYPE: PRT
/ ORGANISM: Streptomyces avermitilis
US-10-156-761-8798

Query Match
Best Local Similarity 100.0%; Score 7; DB 14; Length 516;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 KSVDFIE 231
DB 483 KSVDFIE 489

RESULT 47
US-10-072-012-732
/ Sequence 732, Application US/10072012
/ Publication No. US20040033493A1
/ GENERAL INFORMATION:
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Patutajan, Meera
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangolli, Bsha
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Rastelli, Luca
```

```
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Coleman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Futak, Katarzyna
/ APPLICANT: Grosse, William M.
/ APPLICANT: Alsbrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072,012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 60/266,767
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: 60/267,057
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/266,975
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/267,459
/ PRIOR FILING DATE: 2001-02-08
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1391
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 732
/ LENGTH: 523
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-10-072-012-732

Query Match
Best Local Similarity 100.0%; Score 7; DB 12; Length 523;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 488 RTSSSGT 494
DB 49 RTSSSGT 55

RESULT 48
US-10-282-122A-61404
/ Sequence 61404, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zykind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
```

```

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/151,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61404
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Legionella pneumophila
US-10-282-122A-61404

```

```

Query Match      1.1%; Score 7; DB 12; Length 525;
Best Local Similarity 100.0%; Pred.No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      22 GTEKVL 28
DB      225 GTEKVL 231

```

```

RESULT 49
US-10-361-460-73
; Sequence 73, Application US/10361460
; Publication No. US20030163839A1
; GENERAL INFORMATION:
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Genes Encoding Enzymes for Lignin
; TITLE OF INVENTION: Biosynthesis and Uses Thereof
; FILE REFERENCE: 0709
; CURRENT APPLICATION NUMBER: US/10/361,460
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/057,082
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 73
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Zea mays
US-10-361-460-73

```

```

Query Match      1.1%; Score 7; DB 14; Length 526;
Best Local Similarity 100.0%; Pred.No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      252 ATTAIAY 258
DB      108 ATTAIAY 114

```

```

RESULT 50
US-10-289-757-106
; Sequence 106, Application US/10289757
; Publication No. US20030180751A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Forester, Richard L.
; APPLICANT: Gibson, John Bryan
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: No. US20030180751A1r1s, Geoffrey
; APPLICANT: Glenn, Matthew
; APPLICANT: Saulsbury, Keith Martin
; APPLICANT: Hall, Claire
; TITLE OF INVENTION: Compositions isolated from forage
; TITLE OF INVENTION: Grasses and methods for their use
; FILE REFERENCE: 11000.10610
; CURRENT APPLICATION NUMBER: US/10/289,757
; CURRENT FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/337,703
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Lolium perenne
US-10-289-757-106

```

```

Query Match      1.1%; Score 7; DB 14; Length 532;
Best Local Similarity 100.0%; Pred.No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      252 ATTAIAY 258
DB      109 ATTAIAY 115

```

```

Search completed: June 21, 2004, 13:30:33
Job time : 50 secs

```

This Page Blank (uspto)

Wed Jun 23 06:36:37 2004

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:25:49 ; Search time 22 Seconds
(without alignments)
1560.512 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665

Sequence: 1 MAHEMIGTIVRLVALLE.....LKVGSQSSFSGSMETIEVS 665

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 28498

Minimum DB seq length: 473

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	472	71.0	665	4 US-09-816-494-2	Sequence 2, App11
2	1.2	1.2	530	3 US-08-477-4608-4	Sequence 4, App11
3	1.2	1.2	530	3 US-08-379-516-4	Sequence 4, App11
4	1.2	1.2	530	3 US-09-329-916-4	Sequence 4, App11
5	1.2	1.2	530	3 US-08-485-372A-4	Sequence 4, App11
6	1.2	1.2	530	4 US-09-409-006A-4	Sequence 4, App11
7	1.2	1.2	530	4 US-08-484-681-4	Sequence 4, App11
8	1.2	1.2	530	5 PCT-US93-07422-4	Sequence 4, App11
9	1.2	1.2	532	2 US-08-417-495-6	Sequence 6, App11
10	1.2	1.2	532	2 US-08-384-391B-6	Sequence 6, App11
11	1.2	1.2	532	3 US-09-218-950-6	Sequence 6, App11
12	1.2	1.2	532	5 PCT-US92-01785-6	Sequence 6, App11
13	1.2	1.2	532	5 PCT-US95-00454-6	Sequence 6, App11
14	1.2	1.2	575	2 US-08-417-495-4	Sequence 4, App11
15	1.2	1.2	575	2 US-08-384-391B-4	Sequence 4, App11
16	1.2	1.2	575	2 US-09-218-950-4	Sequence 4, App11
17	1.2	1.2	575	5 PCT-US92-01785-4	Sequence 4, App11
18	1.2	1.2	575	5 PCT-US95-00454-4	Sequence 4, App11
19	1.2	1.2	630	4 US-08-472-888A-6	Sequence 4, App11
20	1.2	1.2	903	1 US-08-021-601-12	Sequence 12, App1
21	1.2	1.2	903	1 US-08-082-849B-12	Sequence 12, App1
22	1.2	1.2	903	5 PCT-US94-01624-12	Sequence 12, App1
23	1.2	1.2	476	4 US-09-134-001C-4458	Sequence 4458, App
24	1.2	1.2	476	4 US-09-543-681A-7036	Sequence 7036, App
25	1.2	1.2	553	4 US-09-252-991A-32621	Sequence 32621, A
26	1.2	1.2	564	4 US-09-437-568A-2	Sequence 2, App11
27	1.2	1.2	569	1 US-08-306-231-3	Sequence 3, App11

ALIGNMENTS

28	7	1.1	585	4	US-09-173-300-7	Sequence 7, App11
29	7	1.1	679	4	US-09-252-991A-28887	Sequence 28887, A
30	7	1.1	687	4	US-09-198-452A-848	Sequence 848, App
31	7	1.1	700	4	US-09-978-594-67	Sequence 67, App1
32	7	1.1	816	2	US-08-533-306A-6	Sequence 6, App11
33	7	1.1	816	2	US-08-742-923A-6	Sequence 6, App11
34	7	1.1	818	4	US-09-134-000C-5599	Sequence 5599, App
35	7	1.1	885	2	US-08-533-306A-4	Sequence 4, App11
36	7	1.1	885	2	US-08-742-922A-4	Sequence 4, App11
37	7	1.1	888	4	US-09-540-235-2916	Sequence 2916, App
38	7	1.1	896	2	US-08-640-389A-10	Sequence 10, App1
39	7	1.1	896	4	US-08-618-957A-10	Sequence 10, App1
40	7	1.1	896	4	US-09-043-816E-13	Sequence 13, App1
41	7	1.1	896	4	US-09-357-914-13	Sequence 33, App1
42	7	1.1	896	4	US-08-780-562-3	Sequence 3, App11
43	7	1.1	898	2	US-08-693-697-36	Sequence 36, App1
44	7	1.1	898	4	US-08-588-189-3	Sequence 3, App11
45	7	1.1	906	2	US-08-640-389A-9	Sequence 9, App11
46	7	1.1	906	4	US-08-618-957A-9	Sequence 9, App11
47	7	1.1	906	4	US-09-357-914-32	Sequence 32, App1
48	7	1.1	908	2	US-08-693-697-33	Sequence 33, App1
49	7	1.1	908	2	US-08-588-526-3	Sequence 3, App11
50	7	1.1	923	4	US-08-780-562-4	Sequence 4, App11
51	7	1.1	958	2	US-08-640-389A-8	Sequence 8, App11
52	7	1.1	958	4	US-08-618-957A-8	Sequence 8, App11
53	7	1.1	960	1	US-08-355-888A-8	Sequence 8, App11
54	7	1.1	960	2	US-08-588-190-3	Sequence 3, App11
55	7	1.1	960	2	US-08-693-697-8	Sequence 8, App11
56	7	1.1	960	2	US-08-640-389A-3	Sequence 3, App11
57	7	1.1	960	3	US-08-693-696-8	Sequence 8, App11
58	7	1.1	960	4	US-08-618-957A-3	Sequence 3, App11
59	7	1.1	960	4	US-09-357-914-8	Sequence 8, App11
60	7	1.1	1036	4	US-09-252-991A-18349	Sequence 18349, A
61	7	1.1	1051	4	US-09-543-681A-8143	Sequence 8143, App
62	7	1.1	1099	3	US-08-726-21-14	Sequence 11, App1
63	7	1.1	1165	2	US-08-640-389A-11	Sequence 11, App1
64	7	1.1	1165	2	US-08-599-455B-4	Sequence 4, App11
65	7	1.1	1165	3	US-09-093-814-1	Sequence 1, App11
66	7	1.1	1165	3	US-09-069-781B-4	Sequence 4, App11
67	7	1.1	1165	4	US-08-618-957A-11	Sequence 11, App1
68	7	1.1	1165	4	US-08-864-564A-4	Sequence 4, App11
69	7	1.1	1165	4	US-09-137-132-4	Sequence 4, App11
70	7	1.1	1165	4	US-09-094-410-4	Sequence 4, App11
71	7	1.1	1165	4	US-08-708-123D-4	Sequence 4, App11
72	7	1.1	1165	4	US-08-583-153A-4	Sequence 4, App11
73	7	1.1	1165	4	US-08-570-142D-4	Sequence 4, App11
74	7	1.1	1165	4	US-08-780-562-2	Sequence 2, App11
75	7	1.1	1165	4	US-08-638-524B-4	Sequence 4, App11

RESULT 1
US-09-816-494-2
Sequence 2, Application US/09816494
Patent No. 6664089
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
FILE REFERENCE: 10448-030002
CURRENT APPLICATION NUMBER: US/09/816,494
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,858
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 665
TYPE: PRT
ORGANISM: Homo sapiens
US-09-816-494-2

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/329,916
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,460
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-329-916-4

Query Match 1.2%; Score 8; DB 3; Length 530;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 605 DRADSRSS 612
Db 78 DRADSRSS 85

RESULT 5
US-08-485-372A-4
Sequence 4, Application US/08485372A
Patent No. 6187748
GENERAL INFORMATION:
APPLICANT: Beaudry, Gary A.
APPLICANT: Maddon, Paul J.
TITLE OF INVENTION: CD4-GAMMA2 CD4-IgG2 CHIMERAS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,372A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,227
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-II-A

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-485-372A-4

Query Match 1.2%; Score 8; DB 3; Length 530;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 605 DRADSRSS 612
Db 78 DRADSRSS 85

RESULT 6
US-09-409-006A-4
Sequence 4, Application US/09409006A
Patent No. 6342586
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/409,006A
FILING DATE: 29-SEP-1999
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-409-006A-4

Query Match 1.2%; Score 8; DB 4; Length 530;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
|||||
DB 78 DRADSRRS 85

RESULT 7

US-08-484-681-4
; Sequence 4, Application US/08484681
; Patent No. 6451313
; GENERAL INFORMATION:
; APPLICANT: Beaudry, Gary A.
; TITLE OF INVENTION: CD4-GAMMA2 CD4-IGG2 CHIMERAS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 37690-II-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: homo sapien
; CELL TYPE: lymphocyte
US-08-484-681-4

Query Match 1.2%; Score 8; DB 4; Length 530;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
|||||
DB 78 DRADSRRS 85

RESULT 8

PCT-US93-07422-4
; Sequence 4, Application PC/TUS9307422
; GENERAL INFORMATION:
; APPLICANT: Progenics Pharmaceuticals, Inc.
; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
; TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IGG2 IMMUNOCONJUGATES, AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07422
FILING DATE: 19930806
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
PCT-US93-07422-4

Query Match 1.2%; Score 8; DB 5; Length 530;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
|||||
DB 78 DRADSRRS 85

RESULT 9

US-08-417-495-6
; Sequence 6, Application US/08417495
; Patent No. 5843728
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Redirection of Cellular Immunity by
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,495
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/203,866

Chimeras

FILED DATE: US/07/847,566
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-417-495-6

Query Match 1.2%; Score 8; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
DB 78 DRADSRSS 85

RESULT 10

US-08-284-391B-6
Sequence 6, Application US/08284391B
Patent No. 5851828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045

TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-6

Query Match 1.2%; Score 8; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
DB 78 DRADSRSS 85

RESULT 11

US-09-218-950-6
Sequence 6, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-6

Query Match 1.2%; Score 8; DB 3; Length 532;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 605 DRADSRSS 612
|||||||
Db 78 DRADSRSS 85

RESULT 12

PCT-US92-01785-6
; Sequence 6, Application PC/TUS9201785
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
; TITLE OF INVENTION: Chimeras
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/01785
; FILING DATE: 19920306
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/665,961
; FILING DATE: March 7, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/119002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-01785-6

Query Match 1.2%; Score 8; DB 5; Length 532;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 605 DRADSRSS 612
|||||||
Db 78 DRADSRSS 85

RESULT 13

PCT-US95-00454-6
; Sequence 6, Application PC/TUS9500454
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
; TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/1247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00454-6

Query Match 1.2%; Score 8; DB 5; Length 532;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 605 DRADSRSS 612
|||||||
Db 78 DRADSRSS 85

RESULT 14

US-08-417-495-4
; Sequence 4, Application US/08417495
; Patent No. 5843728
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Redirection of Cellular Immunity by
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,495
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,866

FILING DATE:
APPLICATION NUMBER: US/07/847,566
FILING DATE: 07/665,961
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-417-495-4

Query Match 1.2%; Score 8; DB 2; Length 575;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
DB 78 DRADSRRS 85

RESULT 15

US-08-284-391B-4
Sequence 4, Application US/08284391B
Patent No. 5851828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elding LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elding, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045

TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-4

Query Match 1.2%; Score 8; DB 2; Length 575;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
DB 78 DRADSRRS 85

RESULT 16

US-09-218-950-4
Sequence 4, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elding LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elding, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-4

Query Match 1.2%; Score 8; DB 3; Length 575;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
Db 78 DRADSRSS 85

RESULT 17

PCT-US92-01785-4
; Sequence 4, Application PC/TUS9201785
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
; TITLE OF INVENTION: Chimeras
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/01785
; FILING DATE: 19920306
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/665,961
; FILING DATE: March 7, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/119002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-01785-4

Query Match 1.2%; Score 8; DB 5; Length 575;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
Db 78 DRADSRSS 85

RESULT 18

PCT-US95-00454-4
; Sequence 4, Application PC/TUS9500454
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
; TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00454-4

Query Match 1.2%; Score 8; DB 5; Length 575;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
Db 78 DRADSRSS 85

RESULT 19

US-08-472-888A-6
; Sequence 6, Application US/08472888A
; Patent No. 6613746
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Walz, Gerd
; TITLE OF INVENTION: AGP-ANTIBODY FUSION PROTEINS
; TITLE OF INVENTION: AND RELATED MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,888A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/618,314
FILING DATE: 23-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/258001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 630 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-888A-6

Query Match 1.2%; Score 8; DB 4; Length 630;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
DB 78 DRADSRSS 85

RESULT 20
US-08-021-601-12
Sequence 12, Application US/08021601
Patent No. 5591631
GENERAL INFORMATION:
APPLICANT: Leppla, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Nichols, Peter J.
APPLICANT: Aroora, Naveen
APPLICANT: Singh, Yogendra
TITLE OF INVENTION: ANTIRAX TOXIN FUSION PROTEINS AND
TITLE OF INVENTION: RELATED METHODS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Needle & Rosenberg, P.C.
STREET: 133 Carnegie Way, Suite 400
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,601
FILING DATE: 19930212
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Spratl, Gwendolyn D.
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 1414.057
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 903 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-021-601-12

Query Match 1.2%; Score 8; DB 1; Length 903;

Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
DB 778 DRADSRSS 785

RESULT 21
US-08-082-849B-12
Sequence 12, Application US/08082849B
Patent No. 5677274
GENERAL INFORMATION:
APPLICANT: Leppla, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Aroora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
TITLE OF INVENTION: Related Methods
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082,849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-161-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 903 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-082-849B-12

Query Match 1.2%; Score 8; DB 1; Length 903;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRSS 612
DB 778 DRADSRSS 785

RESULT 22
PCT-US94-01624-12
Sequence 12, Application PC/TUS9401624
GENERAL INFORMATION:
APPLICANT: Leppla, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Aroora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.

TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
TITLE OF INVENTION: RELATED METHODS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
STREET: Stewart Street Tower, 20th Floor, One Market
STREET: Plaza
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01624
FILING DATE: June 25, 1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-115
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 903 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01624-12

Query Match 1.2%; Score 8; DB 5; Length 903;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 DRADSRRS 612
DB 778 DRADSRRS 785

RESULT 23
US-09-134-001C-4458
Sequence 4458, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stramm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4458
LENGTH: 476
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4458

Query Match 1.1%; Score 7; DB 4; Length 476;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 376 LVQALSG 382
DB 470 LVQALSG 476

RESULT 24
US-09-543-681A-7036
Sequence 7036, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 7036
LENGTH: 486
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-7036

Query Match 1.1%; Score 7; DB 4; Length 486;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 615 ERSPEK 621
DB 380 ERSPEK 386

RESULT 25
US-09-252-991A-32621
Sequence 32621, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32621
LENGTH: 553
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32621

Query Match 1.1%; Score 7; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 AAGORPV 355
DB 152 AAGORPV 158

RESULT 26
US-09-437-568A-2
Sequence 2, Application US/09437568A
Patent No. 6620603
GENERAL INFORMATION:
APPLICANT: Lambeth, J. David
APPLICANT: Griendling, Kathy
APPLICANT: Lassegue, Bernard
APPLICANT: Arnold, Rebecca S.
APPLICANT: Cheng, Guangjie
TITLE OF INVENTION: No. 6620603el Mitogenic Regulators

FILE REFERENCE: 05501-0103
CURRENT APPLICATION NUMBER: US/09/437,568A
CURRENT FILING DATE: 1999-11-10
NUMBER OF SEQ ID NOS: 61
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 564
TYPE: PRT
ORGANISM: Homo sapiens
US-09-437-568A-2

Query Match 1.1%; Score 7; DB 4; Length 564;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 644 BELGKVG 650
DB 460 BELGKVG 466

RESULT 27

US-08-306-231-3
Sequence 3, Application US/08306231
Patent No. 5643748
GENERAL INFORMATION:
APPLICANT: Snodgrass, H. R.
APPLICANT: Cioffi, Joseph
APPLICANT: Zupancic, Thomas J.
APPLICANT: Shafer, Alan W.
TITLE OF INVENTION: Hu-B1.219, A NOVEL HEMATOPOIETIN
TITLE OF INVENTION: RECEPTOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,231
FILING DATE: 14-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 7225-076
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 569 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-306-231-3

Query Match 1.1%; Score 7; DB 1; Length 569;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPLV 377
DB 56 LEDSPLV 62

RESULT 28

US-09-173-300-7
Sequence 7, Application US/09173300
Patent No. 6451581
GENERAL INFORMATION:
APPLICANT: Falco, Saverio Carl
APPLICANT: Hiltz, William D.
APPLICANT: Kinney, Anthony J.
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Rafalski, J. Antoni
TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
FILE REFERENCE: BB-1126
CURRENT APPLICATION NUMBER: US/09/173,300
CURRENT FILING DATE: 1998-10-15
EARLIER APPLICATION NUMBER: 60/063,423
EARLIER FILING DATE: 1997 October 28
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Microsoft Word Version 7.0A
SEQ ID NO 7
LENGTH: 585
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-173-300-7

Query Match 1.1%; Score 7; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 ASNGCVL 241
DB 577 ASNGCVL 583

RESULT 29

US-09-252-991A-28887
Sequence 28887, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 28887
LENGTH: 679
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28887

Query Match 1.1%; Score 7; DB 4; Length 679;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 FLGQLLD 293
DB 397 FLGQLLD 403

RESULT 30

US-09-198-452A-848
Sequence 848, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Grifflais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection

FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 848
LENGTH: 687
TYPE: PRP
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-848

Query Match 1.1%; Score 7; DB 4; Length 687;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 LIQHSK 74
DB 427 LIQHSK 433

RESULT 31
US-09-976-594-67
Sequence 67, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 67
LENGTH: 700
TYPE: PRP
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6673549 3393396CD1
US-09-976-594-67

Query Match 1.1%; Score 7; DB 4; Length 700;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 CSQKVV 87
DB 479 CSQKVV 485

RESULT 32
US-08-533-306A-6
Sequence 6, Application US/0853306A
Patent No. 5837457
GENERAL INFORMATION:
APPLICANT: Liu, Pu
APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/533,306A
FILING DATE: September 25, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869COB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-533-306A-6

Query Match 1.1%; Score 7; DB 2; Length 816;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 DVASLS 100
DB 187 DVASLS 193

RESULT 33
US-08-742-923A-6
Sequence 6, Application US/08742923A
Patent No. 5869611
GENERAL INFORMATION:
APPLICANT: Liu, Pu
APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,923A
FILING DATE: No. 5869611ember 1, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869DVC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-742-923A-6

Query Match 1.1%; Score 7; DB 2; Length 816;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 DVASLSS 100
|||||
DB 187 DVASLSS 193

RESULT 34
US-09-134-000C-5599
; Sequence 5599, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5599
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5599

Query Match 1.1%; Score 7; DB 4; Length 818;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 510 EDNYHTS 516
|||||
DB 337 EDNYHTS 343

RESULT 35
US-08-533-306A-4
; Sequence 4, Application US/08533306A
; Patent No. 5837457
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESS: P.O. Box 828
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/533,306A
; FILING DATE: September 25, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869COB
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-533-306A-4

Query Match 1.1%; Score 7; DB 2; Length 885;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 DVASLSS 100
|||||
DB 256 DVASLSS 262

RESULT 36
US-08-742-923A-4
; Sequence 4, Application US/08742923A
; Patent No. 5869611
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESS: P.O. Box 828
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,923A
; FILING DATE: NO. 5869611ember 1, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869DVC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600
; TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-742-923A-4

Query Match 1.1%; Score 7; DB 2; Length 885;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 DVASLSS 100
|||||
DB 256 DVASLSS 262

RESULT 37
US-09-540-236-2916

Sequence 2916, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540.236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 2916
LENGTH: 888
TYPE: PRT
ORGANISM: M.catarhalis
US-09-540-236-2916

Query Match 1.1%; Score 7; DB 4; Length 888;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 IYTERLV 16
DB 552 IYTERLV 558

RESULT 38
US-08-640-389A-10
Sequence 10, Application US/08640389A
Patent No. 5912123
GENERAL INFORMATION:
APPLICANT: Snodgrass, H. R.
APPLICANT: Cioffi, Joseph
APPLICANT: Zupancic, Thomas J.
APPLICANT: Shafer, Alan W.
TITLE OF INVENTION: DETECTION OF THE LEPTIN
TITLE OF INVENTION: RECEPTOR IN REPRODUCTIVE ORGANS AND METHODS FOR
TITLE OF INVENTION: REGULATING REPRODUCTIVE BIOLOGY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/640,389A
FILING DATE: 29-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION/DOCKET NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-032
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 896 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-640-389A-10

Query Match 1.1%; Score 7; DB 2; Length 896;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPV 377
DB 169 LEDSPV 175

RESULT 39
US-08-618-957A-10
Sequence 10, Application US/08618957A
Patent No. 6355237
GENERAL INFORMATION:
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Cioffi, Joseph
APPLICANT: Zupancic, Thomas Joel
APPLICANT: Shafer, Alan Wayne
TITLE OF INVENTION: METHODS FOR USING THE OBESE
TITLE OF INVENTION: GENE AND ITS GENE PRODUCT TO STIMULATE HEMATOPOIETIC
TITLE OF INVENTION: DEVELOPMENT
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds LLP
STREET: 1155 Avenue of The Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,957A
FILING DATE: 20-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION/DOCKET NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 008907-0033-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 896 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-618-957A-10

Query Match 1.1%; Score 7; DB 4; Length 896;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPV 377
DB 169 LEDSPV 175

RESULT 40
US-09-043-816E-13
Sequence 13, Application US/09043816E
Patent No. 6414128
GENERAL INFORMATION:
APPLICANT: Hilton, Douglas J.
APPLICANT: Willson, Tracy
APPLICANT: Nicola, Nicos A.

```
; APPLICANT: Gainsford, Timothy
; APPLICANT: Alexander, Warren S.
; APPLICANT: Metcalf, Donald
; APPLICANT: Ng, Ashley
; TITLE OF INVENTION: A NOVEL HAEMOPOLYMERIN RECEPTOR AND GENETIC SEQUENCES
; FILE REFERENCE: 11268
; CURRENT APPLICATION NUMBER: US/09/043,816E
; CURRENT FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 13
; LENGTH: 896
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (223)
; OTHER INFORMATION: Xaa is unknown or other.
; NAME/KEY: UNSURE
; LOCATION: (687)
; OTHER INFORMATION: Xaa is unknown or other.
; US-09-043-816E-13
```

```
Query Match      1.1%; Score 7; DB 4; Length 896;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      371 LEDSPV 377
        |||||
Db      169 LEDSPV 175
```

```
RESULT 41
US-09-357-914-33
; Sequence 33, Application US/09357914
; Patent No. 6524806
; GENERAL INFORMATION:
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Cioffi, Joseph
; APPLICANT: Zupancic, Thomas J.
; APPLICANT: Shafer, Alan Wayne
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR Hu-B1.219, A
; FILE REFERENCE: 8907-0083-999
; CURRENT APPLICATION NUMBER: US/09/357,914
; CURRENT FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/693,696
; PRIOR FILING DATE: 1996-08-05
; PRIOR APPLICATION NUMBER: US 08/355,888
; PRIOR FILING DATE: 1994-12-14
; PRIOR APPLICATION NUMBER: US 08/306,231
; PRIOR FILING DATE: 1994-09-14
; NUMBER OF SEQ. ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 896
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-357-914-33
```

```
Query Match      1.1%; Score 7; DB 4; Length 896;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      371 LEDSPV 377
        |||||
Db      169 LEDSPV 175
```

```
RESULT 42
US-08-780-562-3
; Sequence 3, Application US/08780562
```

```
; Patent No. 6541604
; GENERAL INFORMATION:
; APPLICANT: Matthews, William
; APPLICANT: Bennett, Brian
; TITLE OF INVENTION: WSX RECEPTOR
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/780,562
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585005
; FILING DATE: 01/08/97
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/
; FILING DATE: 01/08/97
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0986R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 896 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-780-562-3
```

```
Query Match      1.1%; Score 7; DB 4; Length 896;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      371 LEDSPV 377
        |||||
Db      169 LEDSPV 175
```

```
RESULT 43
US-08-693-697-36
; Sequence 36, Application US/08693697
; Patent No. 5869610
; GENERAL INFORMATION:
; APPLICANT: Snodgrass, H. R.
; APPLICANT: Cioffi, Joseph
; APPLICANT: Zupancic, Thomas J.
; APPLICANT: Shafer, Alan W.
; TITLE OF INVENTION: Hu-B1.219, A NOVEL HUMAN HAEMOPOLYMERIN
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Penile & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,697
; FILING DATE: 05-AUG-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0037-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNITE
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-693-697-36

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 898;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPV 377
DB 171 LEDSPV 177

RESULT 44
US-08-588-189-3
; Sequence 3, Application US/08588189
; Patent No. 6451523
; GENERAL INFORMATION:
; APPLICANT: Snodgrass, H.
; APPLICANT: Cioffi, Joseph
; APPLICANT: Zupancic, Thomas
; APPLICANT: Shafer, Alan
; TITLE OF INVENTION: DETECTION OF A LEPTIN RECEPTOR VARIANT
; TITLE OF INVENTION: AND METHODS FOR REGULATING OBESITY
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,189
; FILING DATE: 18-JAN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741
; TELEX: 66141 PENNITE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids

```

```

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-588-189-3

Query Match
Best Local Similarity 100.0%; Score 7; DB 4; Length 898;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPV 377
DB 171 LEDSPV 177

RESULT 45
US-08-640-389A-9
; Sequence 9, Application US/08640389A
; Patent No. 5912123
; GENERAL INFORMATION:
; APPLICANT: Snodgrass, H. R.
; APPLICANT: Cioffi, Joseph
; APPLICANT: Zupancic, Thomas J.
; APPLICANT: Shafer, Alan W.
; TITLE OF INVENTION: DETECTION OF THE LEPTIN
; TITLE OF INVENTION: RECEPTOR IN REPRODUCTIVE ORGANS AND METHODS FOR
; TITLE OF INVENTION: REGULATING REPRODUCTIVE BIOLOGY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/640,389A
; FILING DATE: 29-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 8907-032
; REFERENCE/DOCKET NUMBER: 8907-032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNITE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-640-389A-9

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 906;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPV 377
DB 169 LEDSPV 175

RESULT 46
US-08-618-957A-9
; Sequence 9, Application US/08618957A
; Patent No. 6355237

```

```

; GENERAL INFORMATION:
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Cioffli, Joseph
; APPLICANT: Zupancic, Thomas Joel
; APPLICANT: Shafer, Alan Wayne
; TITLE OF INVENTION: METHODS FOR USING THE OBSE
; TITLE OF INVENTION: GENE AND ITS GENE PRODUCT TO STIMULATE HEMATOPOIETIC
; TITLE OF INVENTION: DEVELOPMENT
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of The Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,957A
; FILING DATE: 20-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 008907-0033-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-618-957A-9

Query Match 1.1%; Score 7; DB 4; Length 906;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPV 377
Db 169 LEDSPV 175

RESULT 47
US-09-357-914-32
; Sequence 32, Application US/09357914
; Patent No. 6524806
; GENERAL INFORMATION:
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Cioffli, Joseph
; APPLICANT: Zupancic, Thomas J.
; APPLICANT: Shafer, Alan Wayne
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR Hu-B1.219, A
; TITLE OF INVENTION: NOVEL HUMAN HEMATOPOIETIN RECEPTOR
; FILE REFERENCE: 8907-0083-999
; CURRENT APPLICATION NUMBER: US/09/357,914
; CURRENT FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/693,696
; PRIOR FILING DATE: 1996-08-05
; PRIOR APPLICATION NUMBER: US 08/355,888
; PRIOR FILING DATE: 1994-12-14
; PRIOR APPLICATION NUMBER: US 08/306,231
```

```

; PRIOR FILING DATE: 1994-09-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 906
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-357-914-32

Query Match 1.1%; Score 7; DB 4; Length 906;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPV 377
Db 169 LEDSPV 175

RESULT 48
US-08-693-697-33
; Sequence 33, Application US/08693697
; Patent No. 5869610
; GENERAL INFORMATION:
; APPLICANT: Snodgrass, H. R.
; APPLICANT: Cioffli, Joseph
; APPLICANT: Zupancic, Thomas J.
; APPLICANT: Shafer, Alan W.
; TITLE OF INVENTION: Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN
; TITLE OF INVENTION: RECEPTOR
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,697
; FILING DATE: 03-AUG-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0037-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 908 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-693-697-33

Query Match 1.1%; Score 7; DB 2; Length 908;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPV 377
Db 171 LEDSPV 177
```

RESULT 49
US-08-588-526-3
; Sequence 3, Application US/08588526
; Patent No. 5882860
; GENERAL INFORMATION:
; APPLICANT: Snodgrass, H.
; APPLICANT: Cioffi, Joseph
; APPLICANT: Zupancic, Thomas
; APPLICANT: Shafer, Alan
; TITLE OF INVENTION: DETECTION OF A LEPTIN RECEPTOR
; TITLE OF INVENTION: VARIANT
; TITLE OF INVENTION: AND METHODS FOR REGULATING OBESITY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,526
; FILING DATE: 18-JAN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 908 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-588-526-3

Query Match 1.1%; Score 7; DB 2; Length 908;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPV 377
Db 171 LEDSPV 177

RESULT 50
US-08-780-562-4
; Sequence 4, Application US/08780562
; Patent No. 6541604
; GENERAL INFORMATION:
; APPLICANT: Matthews, William
; APPLICANT: Bennett, Brian
; TITLE OF INVENTION: MSX RECEPTOR
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/780,562
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585005
; FILING DATE: 01/08/97
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/
; FILING DATE: 01/08/97
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0986R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 923 amino acids
; TYPE: Amino Acid
; TOPOLOGY: linear
; US-08-780-562-4

Query Match 1.1%; Score 7; DB 4; Length 923;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 LEDSPV 377
Db 169 LEDSPV 175

Search completed: June 21, 2004, 13:29:32
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2004, 05:14:42 ; Search time 6973 Seconds

(without alignments)
4133.532 Million cell updates/sec

Title: US-10-029-345a-109

Perfect score: 665
Sequence: 1 MAHEMIGTQIVTERLVALL.....LGRVSGSSFSGSMETIEVS 665

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1263448

Minimum DB seq length: 1419

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-C/cgn2.1/USPTO.spool/US1002345/runat.21062004.164401.29751/app.query.fasta.1.839
-DB=GenEmbl -QFMT=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=75
-DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=50 -MODE=LOCAL -OUTFMT=ptco
-NORMext -HEAPSIZE=500 -MINLEN=1419 -MAXLEN=2000000000
-USER=US1002345.@CNCN_1_1.4290@runat.21062004.164401.29751 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rod.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hcgo_hum.*
40: em_hcgo_mus.*
41: em_hcgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	665	100.0	5450	6	AX482439
2	665	100.0	5450	6	AX482478
3	472	71.0	1935	9	AY038927
4	472	71.0	1998	6	AX260342
5	472	71.0	2102	6	AX713989
6	472	71.0	2102	9	AK055973
7	472	71.0	2732	6	AX180875
8	472	71.0	3059	6	AX278461
9	472	71.0	3104	6	AX405700
10	472	71.0	3332	6	AX441229
11	472	71.0	3496	6	AX441210
12	472	71.0	3521	9	AB052156
13	472	71.0	3544	6	AX260340
14	472	71.0	3566	9	AF506796
15	472	71.0	3766	6	AX374994
16	472	71.0	4790	6	BD171157
17	472	71.0	4790	6	BD183422
18	472	71.0	4790	9	AB051487
19	472	71.0	4790	6	AB051487
20	472	71.0	4790	6	AX099933
21	472	71.0	4790	6	AX921915
22	394	59.2	172206	9	AC007619
23	394	59.2	188344	2	AC131617
24	371	55.8	2807	9	BC031643
25	350	52.6	5111	6	AX482372
26	304	45.7	2071	6	AX921917
27	304	45.7	3284	9	BC042101
28	178	26.8	1916	6	AX835305
29	178	26.8	1916	9	AK098310
30	71	10.7	244605	2	AC097818
31	67	10.1	4943	10	AB052157
32	67	10.1	4943	10	BC059232
33	67	10.1	4975	10	BC057321
34	66	9.9	242590	2	AC133722
35	66	9.9	250782	2	AC128093
36	64	9.6	2756	6	AX482444
37	64	9.6	4026	6	AF345953
38	64	9.6	4992	10	AF345951
39	61	9.2	201474	2	AC126682
40	61	9.2	236589	2	AC118035
41	53	8.0	3861	10	AF345954
42	53	8.0	4827	10	AF345952
43	47	7.1	5047	9	HSB04384
44	27	4.1	247148	2	AC097776
45	22	3.3	3540	3	AK116903
46	20	3.0	2264	9	BC045110
47	20	3.0	2377	6	AR255944
48	20	3.0	2377	9	HSU27193
49	20	3.0	2415	6	A59887
50	20	3.0	2453	6	A59888

QY	321	GIuPvPvAlProAlaValSerGluGlyGlyGlnIuysSerGluuThrProLeuSerProPro	340
Db	1498	GAACCTGTCCTCGCTGCTCAGAGGGGTGACAGAAACCGAGACCCCTCAGTCCACC	1557
QY	341	CysAlaAPSPSerAlaThrSerGluuAlaIaGlyGlnAProValHisProAlaSerVal	360
Db	1558	TGTGCCGACTGTGCTACTCTCAGAGGACAGACAAAGCCCGTGCATCCGCGCAGCTG	1617
QY	361	ProSerValProSerValGlnProSerLeuLeuGlnuAPSerProLeuValGlnAlaLeu	380
Db	1618	CCNAGCGTCCAGCGTGCAGCGCTCGCTGTTAGAGGACAGCGCGCTGTTACAGCGCTC	1677
QY	381	SerGlyLeuHisIaLeuSerAlaAPArgLeuGlnuAPSerAnlyIaLeuIaArgSerPhe	400
Db	1678	AGTGGCGTCGACCTGTCCGACAGACGCGTGGAAAGCACAAATAGACTCAACGTTCCCTTC	1737
QY	401	SerLeuAPPIaLeuSerValSerIyrSerIaSerMetAlaIaSerLeuHisGlyPhe	420
Db	1738	TCTCTGGATATCAATCACTAGTTTCATATTTACGCCACATGGACAGATCTCTTACATGGCTTC	1797
QY	421	SerSerSerGluAPSPAlaLeuGlnuIyrTyrrIuysProSerThrThrLeuAPGlyThrAsn	440
Db	1798	TCCTCATCAGAAAGAGCTTTGGAAATCTACAAACCTTCCACTACTGTGGATGGACCAAC	1855
QY	441	LyseIuCyGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro	460
Db	1858	AAGCTATCCACAGTTCTCCCTGTTCCAGAAACTATGGAGGACAGACTCCGAAACCAAGTCTCT	1917
QY	461	APlyAPGluGlnuIaSerIleProIyValIaLeuGlnThrAlaArgProSerAPSPSerGln	480
Db	1918	GATAAAGAGGAAGCCAGCATCCCAAGAACTGCGACGCCGACGCTTCACAGACGCAG	1977
QY	481	SerIySAArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeu	500
Db	1978	AGCAAGCATTTGCATTCGTGGTCAGAACACAGCAGCATGGCACCGCCCAAGATCCCTTTA	2037
QY	501	SerProLeuHisIaArgSerGlySerValGlnuAPAnTyrrHisSerPheIuPheGly	520
Db	2038	TCTCCACATCGACGAAAGTGGAGCGCTGGAGGACATTTACACACAGCTTCCTTTCCGCG	2097
QY	521	LeuSerThrSerGlnGlnHisIaLeuThrIyrSerAlaGlyLeuGlyIaLeuIyThrHis	540
Db	2098	CTTTCACACAGCAGACAGACCTCAGAAAGTCTGTGGCTTGGGCTTTAAAGGCTGGGAC	2157
QY	541	SerAPPIaLeuAlaProGlnThrSerThrProSerLeuThrSerSerTriTyrrPheAla	560
Db	2158	TGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGACGTGGATTTTGGC	2217
QY	561	ThrGluSerSerHisAPheTyrrSerAlaSerAlaIaIaIyrglyGlySerAlaSerTyrrSer	580
Db	2218	ACAGAGTCTCTCACACTTCTACTCTGCTCAGCCATCTACGGAGGAGAGTCCAGTTACTCT	2277
QY	581	AlaIyrSerCySerSerGlnLeuProThrCyGlyIaAPGlnValTyrrSerValIaArgArg	600
Db	2278	GCTTCAGCTGCAGCAGCTGCCCCACTTGGCGAGACCAAGTCTATCTGTGTGGCGGGCG	2337
QY	601	GlnIuysProSerAPSPArgAlaAPSerIaArgArgSerTriPHisGluGluSerProPheGln	620
Db	2338	CAGAAAGCCAAAGTACAGAGCTGACTCGGGCGGACCTGCATGAAGAAGCCCCCTTTAA	2397
QY	621	LyseGlnPheIySAArgArgSerCyGlnIuMetGluPheGlyGluSerIleMetSerGluAsn	640
Db	2398	AAGCAGTTTAAACGACGAAGCTGCCAAATGGAAATTGGAGAGAGCATCATGTACAGAAAC	2455
QY	641	ArgSerAPGluGluLeuGlyIyValGlySerGlnIuSerSerPheSerGlySerMetGlu	660
Db	2458	AGGTACACGGGAAGACTGGGGGAAAGTGGGCACTCAGTCTTTCGGGACGATGGAA	2517
QY	661	IleIleGluValSer	665
Db	2518	ATCATTTAGAGTCTCC	2532

[illegible]

```

Db      838 GACGTTTCTCACTGTAAGTCTGGTAACTGGAGAAAGCTTCAACTGTGTCACTG 897
Qy      121 Leu1a61y61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 140
Db      898 CTTCAGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 957
Qy      141 Thr1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 160
Db      958 ACTCTATCTCTCACTGTAAGTCTGGTAACTGGAGAAAGCTTCAACTGTGTCACTG 1017
Qy      161 Arg11eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 180
Db      1018 CGAATCTCTCACTGTAAGTCTGGTAACTGGAGAAAGCTTCAACTGTGTCACTG 1077
Qy      181 Glu1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 200
Db      1078 CAGCAGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGG 1137
Qy      201 Ile1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 220
Db      1138 ATCCCGAGTCTCACTGTAAGTCTGGTAACTGGAGAAAGCTTCAACTGTGTCACTG 1197
Qy      221 Pro1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 240
Db      1198 CCGTGGTGGAGAAATCAGATGATTCATTCAGATGATTCATTCAGATGATTCATTCAG 1257
Qy      241 Leu1a1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 260
Db      1258 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1317
Qy      261 Lys1a1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 280
Db      1318 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1377
Qy      281 Ile1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 300
Db      1378 ATATCTCTCACTGTAAGTCTGGTAACTGGAGAAAGCTTCAACTGTGTCACTG 1437
Qy      301 Glu1a1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 320
Db      1438 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1497
Qy      321 Glu1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 340
Db      1498 GAACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1557
Qy      341 Cys1a1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 360
Db      1558 TGTGCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1617
Qy      361 Pro1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 380
Db      1618 CCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1677
Qy      381 Ser1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 400
Db      1678 AGTGGGCTGACCTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1737
Qy      401 Ser1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 420
Db      1738 TCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1797
Qy      421 Ser1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 440
Db      1798 TCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1857
Qy      441 Lys1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 460
Db      1858 AAGCTATGCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1917
Qy      461 Asp1y61yGlu1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 480

```

```

Db      1918 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1977
Qy      481 Ser1y61yGlu1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 500
Db      1978 AGCAAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2037
Qy      501 Ser1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 520
Db      2038 TCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2097
Qy      521 Leu1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 540
Db      2098 CTTTCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2157
Qy      541 Ser1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 560
Db      2158 TCGATATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2217
Qy      561 Thr1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 580
Db      2218 ACGAGTCTCTCACTGTAAGTCTGGTAACTGGAGAAAGCTTCAACTGTGTCACTG 2277
Qy      581 Ala1ySer1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 600
Db      2278 GCTTACAGTCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2337
Qy      601 Glu1y61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 620
Db      2338 CAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2397
Qy      621 Lys1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 640
Db      2398 AAGCAGTCTTAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2457
Qy      641 Arg1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 660
Db      2458 AGTTCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2517
Qy      661 Ile1eua1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61yPhe1a61y 680
Db      2518 ATCATGAGTCTCT 2532

RESULT 3
AY038927      1935 bp      mRNA      linear      PRI 01-AUG-2001
LOCUS      2518
DEFINITION      Homo sapiens truncated MAPK phosphatase 7 (MKP7) mRNA, complete cds, alternatively spliced.
ACCESSION      AY038927
VERSION      AY038927.1
KEYWORDS      GI:15072488
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1935)
AUTHORS      Montpetit,A., Boily,G. and Smette,D.
TITLE      A detailed transcriptional map of the chromosome 12p12 tumor suppressor locus
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 1935)
AUTHORS      Montpetit,A., Boily,G. and Smette,D.
TITLE      Direct Submission
JOURNAL      Submitted (07-JUN-2001) Hemato-Oncology, Hopital Ste-Justine, 3175
Cote-Ste-Catherine, Montreal, Qc H3T 1C5, Canada
FEATURES
source
1..1935
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1935
/gene="MKP7"
28..459
/gene="MKP7"
CDS

```

```

/!note="minor splice variant"
/codon_start=1
/product="truncated MAPK phosphatase 7"
/protein_id="AAK69770.1"
/db_xref="GI:15072489"
/translation="MAHEMIGTQIVTERVALLSESTKYLIDSRPVEYNTSHILE
AININCSKLMKRLQODKYLITELIOHSAHKYDIDCSOKVYVYDQSSQDVASLSDC
FLTVLLGKLEKPSNVSHLLAGDAEMDMLCVKCQGYLSKA"

```

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 0	1935	472	472	0	0	0
Percent Similarity: 100.00%						
Best Local Similarity: 100.00%						
Query Match: 70.98%						

US-10-029-345A-109 (1-665) x AY038927 (1-1935)

```

QY 194 ThCysProlyserProaphhellieProgluSerHisPheleuArgValProValaanaap 213
Db 443 ACCGTGCAAGAGCTGATTTATCCCGAGCTCATTTCCTGCTGCTGCTGGAATGAC 502
QY 214 SerPheCysgluylleuProTyrPleuAspLysSerValaAspPheilegluylala 233
Db 503 AGCTTTGTGAGAAATTTTGGCTGCTGGAACAAATCAGATTTTCATTGGAAGCA 562
QY 234 LysAlaSeranglCysValleuValHisCysleuAlaGlyIleSerArgSerAlaThr 253
Db 563 AAGCCTCCAAATGATGTTTGTAGTCACTGTTTGTGAGTCTCCGCTCCGACAC 622
QY 254 LlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 273
Db 623 ATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682
QY 274 ValIleGluLysArgProThrIleSerProAsnPheAsnPhelugluLysleuAsp 293
Db 683 GTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 742
QY 294 TyrGluLysLysIleLysAsnGlnThrGlyAlaSerGlyProLysSerLysleuLys 313
Db 743 TATGAGAAAGAAATTAAGAACAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTG 802
QY 314 LeuHisleuGluLysProAsnGluProValaValaSerGluGlyGlnLysSer 333
Db 803 CTGCACTGAGAAAGCAATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 862
QY 334 GluThrProleuSerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArg 353
Db 863 GAGAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 922
QY 354 ProValHisProAlaSerValProSerValProSerValGlnProSerleuLysAsp 373
Db 923 CCGGTGCACTCCGCGCAAGCGTCCAGAGGTGCCAGGCGTGCCTGTTAGAGGAC 982
QY 374 SerProleuValGlnAlaLeuSerGlyLeuHisleuSerAlaAspArgleuGluAspSer 393
Db 983 AGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1042
QY 394 AsnLysleuLysArgSerPheSerleuAspIleLysSerValaSerLysSerAlaAsp 413
Db 1043 AATTAAGCTCAAGGCTCTCTCTGATATCAAAATGATTCATATTCAGCACACATG 1102
QY 414 AlaAlaSerleuHisgluPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSer 433
Db 1103 GCAGCATCTTACATGCTCTCTCTCATCAGAAAGCTTGAATATCAAAACCTTCC 1162
QY 434 ThrThrleuAspGlyThrAsnLysleuGlnPheSerProValGlnleuSerGlu 453
Db 1163 ACTACTCTGATGGAGCAAGCAAGCTATGCAATGCTCTCTCTCTCTCTCTCTCTCTCT 1222
QY 454 GluThrProgluThrSerProAspLysGluGluAlaSerleuProLysLysleuGlnThr 473

```

```

Db 1223 CAGACTCCCGAAACCAAGTCTATTAAGAGAAAGCAGCATCCCAAGAAAGTGCAGACC 1282
QY 474 AlaArgProSerAspSerGlnSerLysArgleuHisSerValArgThrSerSerGly 493
Db 1283 GCAGAGCTCTTCCAGACGCGCAAGCAAGCATTTGCTGCTGCAAGCAAGCAGCAGTGGC 1342
QY 494 ThrAlaArgSerleuSerleuSerProleuHisArgSerGlySerValGluAspAspThr 513
Db 1343 ACCGCGCAGAGTCCCTTTATCTCCATGCTGATCGAAGTGGAGCGTGGAGCAATTC 1402
QY 514 HisThrSerPheleuPhegluSerThrSerGlnGlnHisleuThrLysSerAlaGly 533
Db 1403 CACACAGCTTCTCTTTGGCTTCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1462
QY 534 LeuGlyLysleuGlyTyrPheIleSerAspIleAlaProGlnThrSerThrProSerleu 553
Db 1463 CTGGGCTTAAAGGCTGCGCACTCGATATCTGGCCCCCAGACCTCTACCTCTCTCTG 1522
QY 554 ThrSerSerTyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyr 573
Db 1523 ACCAGCAGCTGATATTTGCAAGAGTCTTCACTTACTTCTGCTGCTGCTGCTGCTGCTG 1582
QY 574 GlyLysSerAlaSerTyrSerAlaTyrSerCysSerGlnleuProThrCysgluAspGln 593
Db 1583 GAGAGCAGTCCAGATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1642
QY 594 ValTyrSerValArgArgArgGlnLysProSerAspArgAlaAspSerArgSerTyr 613
Db 1643 GTCTATTTCTGCTGCGCAGCGCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTG 1702
QY 614 HisGluLysSerProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGly 633
Db 1703 CATGAAGAGAGCCCTTTGAAAGCAAGTTAAACGAGAAAGCTGCAAAATGAAATTTGA 1762
QY 634 GluSerIleMetSerGluAsnArgSerArgGluGluLysGlyLysValGlySerGlnSer 653
Db 1763 GAGAGCATATGTCAGAAACAGGTCAACGGAAGAGTGGGGAAGTGGGCACTGCT 1822
QY 654 SerPheSerGlySerMetGluIleleuValSer 665
Db 1823 AGCTTTTCGCGCAGCATGGAATCATTTGAGTCTCC 1858

```

RESULT 4
AX260342
LOCUS AX260342 1998 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 3 from Patent WO0173059.
ACCESSION AX260342
VERSION AX260342.1 GI:16509305
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homosapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Meyers, R.A.
38692 and 21117: dual specificity phosphatase molecules and uses
therefor
Patent: WO 0173059-A 3 04-OCT-2001;
JOURNAL Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..1998
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:
Score: 0	1998	472	472	0	0
Percent Similarity: 99.40%					
Best Local Similarity: 99.40%					
Query Match: 70.98%					

DB: 6 Gaps: 0
US-10-029-345a-109 (1-665) x AX260342 (1-1998)
QY 1 MetAlHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
Db 1 ATGGCCCATGATGATGATTGAACTCAAAATGTTACTAGAGAGGTGGGTGCTGTGGAA 60
QY 21 SerGlyThrGluLeuValLeuLeuIleAspSerArgProPheValGluIlyrAntThrSer 40
Db 61 AGTGAACCGAAAGAGCTGTAAATGATAGCCGGCATTTGTGAATAACAATACATCC 120
QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLeuLeuMetLysArgArgLeuGlnGln 60
Db 121 CACATTTTGGAAACCTTAATATCACTGCTCCAACTTATGAAAGGAGTTGGCAACG 180
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 181 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAACATTAAGGTGACATTGAT 240
QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 241 TGAGTACAGAGGTGTAGTTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 300
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db 301 GACTGTTTCTCATCTGACTTCTGGGTAACCTGGAAGAAGCTTCAACTGTGTACCTG 360
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
Db 361 CTTCGAGGTGGGTTCCTGAGTCTCTGTTGTTCCCTGGCTCTGTGAAGGAAATCC 420
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 421 ACTTAGTCCCTACCTGCAATTCCTCAGCCTGCTTACCTGTGCCAACATTTGGGCCAAC 480
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeuIle 180
Db 481 CGAATTCCTCCCAATCTTATATCTGGCTGCCAGGAGATGCTCTCAACAAGAAGCTGAT- 539
QY 181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProLysProAsp 200
Db 540 GGAGCGAATGGATGTGTTATGTGTTAAATGCCACAA-TACCTGTCCAAAGCCTGACT 598
QY 200 heIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleL 220
Db 599 TTATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATACAGCTTTTGTGAAGAAATTT 658
QY 220 euProTribLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCys 240
Db 659 TCCCGGTGGTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCATGATGTG 718
QY 240 alleuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 260
Db 719 TTCTATGTGACCTGTTAGCTGGGATCTCCGCTCCGCCACCAATCGCTATCGCTATCA 778
QY 260 eLysArgMetLysMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgPro 280
Db 779 TGAAGAGATGAGATGCTTTAGATGAAGCTTACGATTGTGAAAGAAAAAGCCCTA 838
QY 280 hiIleSerProAsnPheAsnLeuGlyGlnLeuLeuAspTyrGluLysLysIleLys 300
Db 839 CTATATCTCCAAACTTCATTTTCTGGGCCAACTCCTGAGCTATGAGAAAGATTAGA 898
QY 300 snGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlnLysPro 320
Db 899 ACCAGACTGAGAGATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTCGAGAAAGCCA 958
QY 320 snGlnProValProAlaValSerGlnGlyGlnLysSerGlnLysProLysSerProp 340
Db 959 ATGAACCTTCTCCTGTGTCTCAGAGGGTGAGACAGAAAGCAAGAGCCCTCAGTCCAC 1018
QY 340 rCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSer 360

Db 1019 CCTGTGCCACTCTCTACTACAGAGCAGACAGCAAAAGCCCTGATCCCGCAGAGG 1078
QY 360 aIProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAla 380
Db 1079 TCCCCAGAGGTGCCAGCTGACGCGCTGCTGTTAGAGACAGCCCGCTGTACAGGCGC 1138
QY 380 euSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSer 400
Db 1139 TCAGTGGCTGCACCTGTCCGAGACAGCTGGAAGACAGCAATAGCTCAAGCTCTCT 1198
QY 400 heSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGly 420
Db 1199 TCTCTCGGATATCAATACAGTTTACTATTTCAGCAGCATGSCACATCTTACTATGCT 1258
QY 420 heSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThr 440
Db 1259 TCTCTCATACAGAAATCTTTGGAATCTACAAACCTTCCACTACTCTGGATGGAGCA 1318
QY 440 snLysLeuGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSer 460
Db 1319 ACAAGCTATGCCAGTTCCTCCCTGTTCAGAACTATCGAGAGACTCCCGAAACAGTC 1378
QY 460 rOAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSer 480
Db 1379 CTGATTAAGAGAGAACCCAGCATCCCAAGAAAGCTGCAACCGCCAGCCTTCAGACAGC 1438
QY 480 InSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
Db 1439 AGAGCAACCGATGTGATTCGGTCAGAACACACAGACAGAGGAGCCGCCAGAGGTCCCT 1498
QY 500 euSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPhe 520
Db 1499 TATCTCCATGCTGATGAGTGGAGCTGAGAGACAAATTACACACAGCTTCTTTTCG 1558
QY 520 IlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTrp 540
Db 1559 GCCTTTCACACAGCAGCAGACGCTCACGAAGTCTGTGGCTGGCTTAAAGGCTGGC 1618
QY 540 iAspSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTyrPhe 560
Db 1619 ACTCGATATCTTGGCCCCCAGACCTTACCTCCCTTCCACAGCAGCTGTGATTTTG 1678
QY 560 iAthrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyr 580
Db 1679 CCAGAGTCTTCACACTTCTACTTGTGCTTACGCCATTTACGAGAGCGATGCCATTTCT 1738
QY 580 eAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
Db 1739 CTGCTTACAGCTGACAGCAGCTGCCCACTTCCGAGAACCAAGTCTATTCTGTGCGCAGC 1798
QY 600 rGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSerProPhe 620
Db 1799 GGCAAAACCAAGTACAGACTGACTCGCGCGAGCTGGCAATGAAAGAGCCCTTTG 1858
QY 620 IuLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGlu 640
Db 1859 AAAAGCATTTAAACGCAAGAGCTGCCAAATGGAATTTGGAGAGCATCTGTCAAGGA 1918
QY 640 snArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMet 660
Db 1919 ACAAGTCACGGGAAGAGCTGGGGAAGTGGCAGTCAAGTCTTCTTTCCGCGAGCATAG 1978
QY 660 IuIleIleGluValSer 665
Db 1979 AAATCATTTGAGTCTCC 1995
RESULT 5
AX713989
LOCUS AX713989 2102 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 673 from Patent EP1293569.
ACCSSION AX713989
VERSION AX713989.1 GI:29888917

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 Isogai, T., Sugiyama, T., Otsuki, T., Makamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Negahari, K. and Masuko, Y.
TITLE Full-length cDNAs
JOURNAL Patent: EP 1293569-A 673 19-MAR-2003;
Helix Research Institute (JP) ; Research Association for Biotechnology (JP)
FEATURES
source location/Qualifiers
1..2102 /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2102
Score: 472.00 Matches: 662
Percent Similarity: 99.10% Conservative: 0
Best Local Similarity: 99.10% Mismatches: 3
Query Match: 70.98% Indels: 6
DB: Gaps: 0
US-10-029-345A-109 (1-665) x AX713989 (1-2102)
QY 1 MetAlaHiegluMetilegluThrglnileValThrgluArgleuValAlaleuEnglu 20
Db 56 ATGGCCCAAGATGATGATTGAACTCAATATTGTTACTGAGAGGTTGGCTCTGCGGAA 115
QY 21 SerGlyThrgluThyValleuLeuileuLeuSerArgProPheValGluTyraenThser 40
Db 116 AGTGAACCGGAAAGGCTGCTAATTGATGACCGGCAATTGTGGATGCAATATACATC 175
QY 41 HieilileuEngluAlaleuLeuileuLeuSerArgleuLeuMetlyArgArgleuEnglu 60
Db 176 CACATTTTGGACCCATTAATATCACTGCTCCAGCTTATGAAGCGAAGGTTGCAACG 235
QY 61 AspIleValleuLeuThrgluLeuileuLeuSerAlaleuHieilileuValleuLeuSer 80
Db 236 GACAAAGCTGTTATACAGAGCTCACTCCAGCATTCAGCAACATTAAGTTGACATTGAT 295
QY 81 CysSerGlnlyValValValValTyraPglInserSerGlnAspValAlaleuLeuSer 100
Db 296 TGCAGTCAGAGGTTGATTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 355
QY 101 AspCysPheleuThrValleuLeuEnglyValleuEnglyValleuSerPheAsnSerValHleu 120
Db 356 GACTGTTTCTCACTGACTTCTGGGTAAACTGAGAGAGCTTCAACTCTGTTCACTG 415
QY 121 LeuAlaGlyValPheleuLeuPheSerArgCysPheProGlyLeuEnglyValleuSer 140
Db 416 CTTCAGAGGTTGGTTGCTGAGTCTCTCTGTTTCTCTGCTGCTGAGGAAATATCC 475
QY 141 ThrIleuValProThrCysIleSerGlnProCysLeuProValAlaleuIleGlyProThr 160
Db 476 ACTCTAGTCCCTACCTGCAATTTCTCAGCCCTTGCTTACCTGCTCCAACTGGGCCAAC 535
QY 161 Arg-11leuProAsnleuTyrlenglyCysGlnArgAspValleuAsnlyValleuLeu 180
Db 536 CT-AATTTCTCCCAATCTTATCTTGCTCCAGCGAGATGCTCTCAACAGAGCTGAT 594
QY 180 e-GlnGlnAsnGlyIleGlyTyraValleuAsnAlaleuTyra-1ThrCysProIysProAsp 199
Db 595 -GACGAGAAATGGGATTTGATTATGTTAAATGCAGCAA-1TACCTGCTCAAGGCTGAC 652
QY 200 PheIleProGluSerHisPheleuArgValProValAsnAspSerPheCysGluTyrlle 219

Db 653 TTTATCCCGAGTCAATTTCTGCGGTGCTGATGATGAGACCTTTTGTGAGAAAT 712
QY 220 LeuProThrLeuAspIleValleuPheileuGlyValAlaleuValleuSerArgleuGly 239
Db 713 TTGGCGGTGGTGGCAAAATCAATGATGATTTTCAATTGAGAAAGCAAAACCTTCATGGATG 772
QY 240 ValIleuValHieCysleuAlaleuGlyIleSerArgSerAlaleuThrIleAlaleuTyrlle 259
Db 773 GTTCTAGTCACTGTTTACTGGATCTCCGCTCCGCCACCACTGCTTACGCTTACATC 832
QY 260 MetIleArgMetAspMetSerleuAspGluAlaTyraArgPheValleuGlyValArgPro 279
Db 833 ATGAAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACCT 892
QY 280 ThrIleSerProAsnPheAsnPhelenglyGlnleuLeuAspTyrlGlyValleuIleVal 299
Db 893 ACTAATATCTCAAACTCAATTTCTGGGCACTCTGAGCTTATGAGAAGATTAAG 952
QY 300 AanglnThrglyValAserGlyProIysSerlyValleuValleuHieilileuEnglyVal 319
Db 953 AACCACTGAGACATCAGGCGCAAGCAAACTCACTGCTGCACTGAGAAAGCA 1012
QY 320 AangluProValProAlaValSerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 339
Db 1013 AATGAACCTGCTCCCTGCTCTCAGAGGGTGAAGAGAAAGGAGAGCGCCCTCACTCA 1072
QY 340 ProCysAlaAspSerAlaThrSerGlyAlaAlaGlyGlnArgProValHisProAlaSer 359
Db 1073 CCTGTGCGAGCTGCTCACTCACTGAGGAGAGAGCAAGAGCCCTGATCCCGCACG 1132
QY 360 ValProSerValProSerValGlnProSerleuLeuGluAspSerProleuValGlnAla 379
Db 1133 GTGCCAGCTGCGCCAGGCTGAGGCTGAGGCTGCTTGAAGAGAGAGGCGCTGATACAGGG 1192
QY 380 LeuSerGlyLeuHieilileuSerAlaAspArgleuGluAspSerAsnlyValleuArgSer 399
Db 1193 CTCAGTGGGCTGACCTGCTCCGAGAGAGCTGAGAGAGCAAGCAATTAAGCTCAAGCTTCC 1252
QY 400 PheSerleuAspIleIleSerValSerTyraSerAlaSerMetAlaAlaleuHieilileu 419
Db 1253 TTCTCTGAGATATCAAAATCAATTTCTATATTCAGGACGATGACACATCTTAACTATG 1312
QY 420 PheSerSerGlyAspAlaleuGluTyrlTyraPglInserThrThreAspGlyTyra 439
Db 1313 TTCTCTCACTCAAGAGATCTTTGAAATCTCAAACTTCCATCTCTGATGGAGAC 1372
QY 440 AsnlyValleuCysGlnPheSerProValGlnleuLeuSerGlnGlnThrProGluThrSer 459
Db 1373 AACCAAGCTATGCAAGTTCTCCCTGTTCAAGAACTATCGAGAGACTCCCGAAACCAAT 1432
QY 460 ProAspIleGluGluAlaleuIleProIlyValleuGlnThrAlaArgProSerAspSer 479
Db 1433 CCTGATTAAGAGAAAGCCAGCAATCCCAAGAGCTGCAACAGCTTCCAGACAGC 1492
QY 480 GlnSerlyArgleuHieilileuValArgThrSerSerSerGlyThrAlaGlnArgSerleu 499
Db 1493 CAGAGCAAGCATTTGATTTGCTGATCAACCAAGCAAGATGGAGCCGCCAGAGTCTCT 1552
QY 500 LeuSerProleuHieilileuSerGlySerValGluAspAsnTyraHieilileuPheleu 519
Db 1553 TTTATCCCACTGATGAGATGGAGCGTGAAGCAATTAACACCAAGCTTCTCTTTC 1612
QY 520 GlyLeuSerThrSerGlnGlnHieilileuThrylValleuGlyValleuGlyValleu 539
Db 1613 GGCCTTTCCACAGCAGCAGCACTCAAGAGTCTGCTGAGGCTTGAAGGCTGAG 1672
QY 540 HisSerAspIleleuAlaProGlnThrSerThrProSerleuThrSerSerTyrlTyraPhe 559
Db 1673 CACTGGATATCTTGGCCCCCAAGACTTACCTTCCCTTACCAAGAGCTGGATTTT 1732
QY 560 AlaThrgluSerSerHisPheTyraSerAlaleuTyraIleTyrlGlyValleuSerTyra 579
Db 1733 GCACAGAGATCTCACTTACTGCTGCTGAGGATCTTACGAGGAGCTGCAAGTTAC 1792

QY 580 SerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArg 599
Db 1793 TCTGCTACAGCGCGCCAGCGCTGCCACCTTGCGGAGCAAGTCTATTCTGTGGCCAG 1852
QY 600 ATGGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGlnGlnSerProPhe 619
Db 1853 CGGCGAAGCCCAAGTACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGCCCTTT 1912
QY 620 GlnLysGlnPheLysArgArgSerCysGlnMetGlnPheGlnGlnSerIleMetSerGln 639
Db 1913 GAAAGACAGTTTAAAGCAGAGCTCCAAATGAAATTTGAGAGAGCATATGTCAGAG 1972
QY 640 AanaSerArgGlnGlnLeuGlnLysValGlySerGlnSerSerPheSerGlySerMet 659
Db 1973 AACAGGTTCACGGGAAAGCTGGGGAAGTGGGAGTCAGCTTTCCTTTCGGGAGCAGATG 2032
QY 660 GlnIleIleGlnValSer 665
Db 2033 GAAATCATTTAGGCTCTCC 2050
RESULT 6
AK055973
LOCUS
DEFINITION Homo sapiens cDNA FLJ1411 fls. clone NT2NE2000214, moderately
similar to DUAL SPECIFICITY PROTEIN PHOSPHATASE 8 (EC 3.1.3.48).
ACCESSION AK055973
VERSION AK055973.1 GI:16550835
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE
AUTHORS 1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,
Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, D., Isono, Y.,
Kawai-Hiro, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsumoto, K., Nakamura, Y., Sekine, M., Kinuchi, H., Kanda, K.,
Magatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagaharti, K., Masuno, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2102)
TITLE Isogai, T., Otsuki, T. and Sugiyama, T.
JOURNAL Direct Submision
REFERENCE Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.
FEATURES
source
1. location/Qualifiers
1. 2102
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2NE2000214"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2NE2"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuron
after the differentiation of NT2 neuronal precursor
cells."
CDS
56..2053
/note="unnamed protein product"
/codon_start=1

/protein_id="BAB71060.1"
/db_xref="GI:16550835"
/translation="MAHEMTCQIVTERVALLSEGTREKVLIDSRPREVYNTSHILE
AINTNSKLMKRRLQDDRVLTTELIGHSARKVIDDCQKVVYVDQSSQDPAASSDC
FLTVLGLKESFNSVHLIAGFAFESFGFLGEGKSLVPTCTISQCLPVANGPT
LILPLVLYGCDRDVINKELMOONGIGVLYNANSPKPDFIPESHFLRPVNDSEK
ILPMLDSEVDFIERAKANGCVLYCLAGISRSATIALAYIMKMDMLDIAFRVKE
KRPTSPNFNFGOLLDEKKIKNOTGASGSKLKLHLKHPNEPVPVSEGGQKSE
TLPSPCDASATSEPAQGRPVHPASVPSPVSEVDSPLVVALSGHLISADILED
SNLKRSPSLDIKVSYSASMAASLHFGSSSEDLLEYEKESTTLDGTRKLCQFSPVOE
LSEQTPETSPDKEASIPKLTQTPASPSQSKRLHSVYTSSSGLDQRLSLPLHSGS
VEDNYHVSFLGLSTSQOHLTKSAGLGLKMHMSIILAQTSPTSLTSSVYATBESHF
YSASAIYGSASAYASCSQPLTCGDVYSVRRQKPSDRADSRSRMHEESPFEKQFK
RRSCQMERGESIMSENRREELGKVGSGSPSSGWEIIEVS"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2102
Score: 472.00 Matches: 662
Percent Similarity: 99.10% Conservative: 0
Best Local Similarity: 99.10% Mismatches: 3
Query Match: 70.96% Indels: 6
DB: 9 Gaps: 0
US-10-029-345A-109 (1-665) x AK055973 (1-2102)
QY 1 MetAlaHisGlnMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGln 20
Db 56 ATGGCCAGTGAAGATGATTGGAACCTAAATTTGATGAGAGGTGTGCTCTGCTGGAA 115
QY 21 SerGlyThrGlnLysValLeuLeuLeuIleAspSerArgPropheValGluTyrAsnThrSer 40
Db 116 AGTGAAGGAAAGAAAGTCTGCTAATTGATGACGGCCATTTGGAAATACATCATCC 175
QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGln 60
Db 176 CACATTTTGAAGACCATTAATATCAACCTCCACAGCTTAAGAAGCGAAGTTGCAACAG 235
QY 61 AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 236 GACCAAGGTATTAATACAGAGCTCATCCAGCATTCAGGAAACATTAAGCTTGACATTAT 295
QY 81 CysSerGlnLysValAlaValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 296 TGCACTCAGAGAGGTGTGATTAGATCAAGAGCTCCAAAGTGTGCTCTCTCTTCA 355
QY 101 AspCysPheLeuThrValLeuLeuGlnLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db 356 GACTGTTTTCACAGTACTTCTGGGTAACTGGAGAAAGAGCTTCAACTGTTTCACTG 415
QY 121 LeuAlaGlyPhePheAlaGlnPheSerArgCysPheProGlyLeuCysGlnGlyLysSer 140
Db 416 CTGCAAGGTGGTGTGAGAGTCTCTGTTGTTCCCTGGCTCTGTGAAGAAATCC 475
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 476 ACTCAAGCCCTACCTGATTTCTGAGCTTGCTTACTGTTGCCAATTTGGGCAACC 535
QY 161 Arg-IleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGlnLeu 180
Db 536 CT-AATTTCTCCATCTTTATCTTGGCTGCGAGGAGATGCTTCCAAAGAGACTAT 594
QY 180 e-GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProLysProAsp 199
Db 595 -GCAGCAAGATGGGATGTATGTATGTTAAAGCCAGAA-TACCTGTCCAAAGCTTAC 652
QY 200 PheIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIle 219
Db 653 TTTATCCCGAGTCACTTCTGCTGCTGCTGATGATGACAGCTTTTGTGAGAAATTT 712
QY 220 LeuProTyrPheLysPheSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCys 239
Db 713 TTGCCGTGTGTGACAAATCAGTATTCATTGAGAAAGCAAAAGCTTCCAAATGAGATG 772

QY 240 ValIeuValHisGlyLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 259
 DB 773 GTTCTAGTGCACGCTTTAGCTGGAGATTCCTCCGCTCCGACACATCGCTTACGCTCATTAC 832
 QY 260 MetIleArgMetAspMetSerIleuAspGluAlaTyrArgPheValIleuGlyIleuArgPro 279
 DB 833 ATGAAAGATGATGATGCTTTAGATGAAGCTTTACAGATTTGTGAAGAAAAGAAAGCCCT 892
 QY 280 ThrIleSerProAsnProAsnProIleuGlyIleuLeuAspTyrGlyIleuValIleuArg 299
 DB 893 ACTATATCTCCAACTTCATTTCTGGGGCAACTCTGACCTATGAGAAAGATTAAAG 952
 QY 300 AsnGlyThrGlyIleuSerGlyProIleuSerIleuLeuValIleuGlyIleuArgPro 319
 DB 953 AACCAAGCTGAGAGCTAGAGCCCAAGCAAACTCAAGCTGCTGACCTGGAGAAAGCCA 1012
 QY 320 AsnGluProValProAlaValSerGluGlyIleuValSerGluThrProIleuSerPro 339
 DB 1013 AATGAACCTGCTCCCTGCTGCTCAGAGGGTGAACAGAAAGGAGAGCCCTCAGTCCA 1072
 QY 340 ProCysAlaAspSerAlaThrSerGluAlaAlaGlyIleuArgProValHisProAlaSer 359
 DB 1073 CCTGTGCCGACTCTGCTACCTCAGAGGAGAGCAAAAGCCCGTGCATCCCGGACAGC 1132
 QY 360 ValProSerValProSerValGlnProSerIleuGluAspSerProIleuValGlnAla 379
 DB 1133 GTGCCAGCTGCGCCAGCGTGACGCGCTGCTGTAGAGACAGCCCGCTGGTACAGGCGC 1192
 QY 380 LeuSerGlyIleuHisLeuSerAlaAspArgLeuGluAspSerAsnIleuValSerArgSer 399
 DB 1193 CTCAGTGGGCTGACCTGCTCGCAGACAGGCTGGAAGACAGCATTAAGCTCAAGCCTTCC 1252
 QY 400 PheSerIleuAspIleuValSerValSerTyrSerAlaSerMetAlaAlaSerIleuHisGly 419
 DB 1253 TTCTCTGTGATTCAAATCAGTTTCAATTCAGCAGCATGCGAGATCTTACATGAGC 1312
 QY 420 PheSerSerSerGluAspAlaLeuGluTyrTyrTyrProSerThrThrIleuAspGlyTyr 439
 DB 1313 TTCTCTCATCAGAAAGTGTGTAATCTTCAAACTTCCACTCTGAGTGGAGACC 1372
 QY 440 AsnIleuLeuCysGlnPheSerProValGlnIleuSerGluGlnThrProGluThrSer 459
 DB 1373 AACAACTATGCAAGTTCTCCCTGTTACGAACTATCGAGAGAGACTCCCGAAACAGT 1432
 QY 460 ProAspIleuGluGluAlaSerIleuProIleuValSerIleuThrAlaArgProSerAspSer 479
 DB 1433 CCTGATAGAGAGAGAGCAGCATCCCAAGAAAGCTCAGACTCCAGGCTTCAGACAGC 1492
 QY 480 GlnSerIleuArgIleuHisSerValArgThrSerSerSerGlyTyrThrAlaGlnArgSerIleu 499
 DB 1493 CAAAGCAAGGATTTGATTCGACAAACAGAGAGAGTGGACCCCGCAGAGGCTCTT 1552
 QY 500 LeuSerProIleuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPhe 519
 DB 1553 TTATCTCCACTGATGAGAGTGGAGCGTGGAGAGACATTAACACACAGCTTCTTTC 1612
 QY 520 GlyLeuSerThrSerGlnGlnHisLeuThrIleSerAlaGlyIleuGlyIleuValIleuArg 539
 DB 1613 GGCCTTTCCACAGCAGCAGCAGCAGCTCAGAAAGTCTGAGCTGGAGCTTAAAGGAGCTG 1672
 QY 540 HisSerAspIleuLeuAlaProGlnThrSerThrProSerIleuThrSerSerTyrPhe 559
 DB 1673 CACTCGAATATCTTGGCCCCCAGACCTTCACTCCCTGACAGCAGCTGGATTTT 1732
 QY 560 AlaThrGlySerSerHisPheTyrSerAlaSerAlaIleTyrGlyIleuSerAlaSerTyr 579
 DB 1733 GCCACAGAGCTCTCACACTTCTGCTCAGCAGCATTCAGAGAGGAGCTGCACTTAC 1792
 QY 580 SerAlaTyrSerCysSerGlnIleuProThrCysGlyAspGlnValTyrSerValArgArg 599
 DB 1793 TCTGCTACAGCTGACAGCCAGCTGCGCACTTGGGAGAGCAAAAGTCTATCTGTGGCAGG 1852

QY 600 ArgGlnIleuProSerAspArgAlaAspSerAlaArgSerTyrHisGluIleuSerProPhe 619
 DB 1853 CCGCAGAACCCAGATGACAGAGCTGACTCGCGCGAGAGCTGGCATTAAGAGACCCCTTT 1912
 QY 620 GluIleuGlnPheIleuArgArgSerCysGlnMetGluPheGlyIleuSerIleuMetSerGlu 639
 DB 1913 GAAAGCAGCTTTAAAGCAGAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCAGAG 1972
 QY 640 AsnArgSerArgGluIleuGlyIleuValGlyIleuSerGlnSerSerPheSerGlySerMet 659
 DB 1973 AACAGGTACCGGAAAGAGCTGGGAAAGTGGCAGTCAAGCTTTTGGGGCAGCATG 2032
 QY 660 GluIleuIleuGluValSer 665
 DB 2033 GAAATCTTGAAGTCTCC 2050
 RESULT 7
 AX180875 2732 bp DNA linear PAT 06-AUG-2001
 LOCUS AX180875
 DEFINITION Sequence 2 from Patent WO0146394.
 ACCESSION AX180875
 VERSION AX180875.1 GI:15132703
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 Ploewman, G.D., Martinez, R., Whyte, D., Manning, G., Sudarsanam, S., Hill, R.J., and Flanagan, P.
 TITLE Mammalian protein phosphatases
 JOURNAL Patent: WO 0146394-A 2 28-JUN-2001; Sugen, Inc. (US)
 FEATURES
 source location/Qualifiers
 1..2732
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2732
 Score: 472.00 Matches: 663
 Percent Similarity: 99.40% Conservative: 0
 Best Local Similarity: 99.40% Mismatches: 2
 Query Match: 70.98% Indels: 4
 DB: 6 Gaps: 0
 US-10-029-345A-109 (1-665) x AX180875 (1-2732)
 QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 DB 538 ATGGCCCATGAGATGATTTGAACTCAAAATTTTACTGAGAGAGTGGTGGCTGCTGGAGAA 597
 QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 DB 598 AGTGAACGGAAGAAAGTGTCTTAATGTATACCGGCAATTTGTGAATACAAATACATCC 657
 QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIleuMetIleuValArgArgLeuGlnGln 60
 DB 658 CACATTTTGAAGCCATTATATCAATGCTCTCAAGCTTATGAGAGAAAGTTGCAACAG 717
 QY 61 AspIleuValLeuIleThrGluLeuIleGlnHisSerAlaIleuHisIleuValAspIleuArg 80
 DB 718 GACAAAGTGTAAATTAACAGACTCATCCAGATTTCAGCGAAACATTAAGGTTGACATTGAT 777
 QY 81 CysSerGlnIleuValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB 778 TGCAGTCGAAGAGTTGTAGTTTACATCAAAAGCTCCCAAGAGTGGCTCTCTCTTCA 837
 QY 101 AspCysPheLeuThrValLeuLeuGlyIleuValSerPheAsnSerValHisIleu 120
 DB 838 GACTGTTTCTCAGTGTACTTCTGGGTAAACTGGAGAAAGAGCTTCACTCTTCACTG 897

ORIGIN

TPUSPCADSAISSEAQRVHPASVSPVQPSLLEDSPLVQALSGHLGADRIED
 SNLKRSFSLDIKSVSASMASLHGFSSDEALRYRPSLTIDTNLCOFSPVOE
 LSEQETSPDKESAI PKKLTARASDSKSLHSPSSGTSADRSLSLSPHRSQS
 VEDNHTSPFLGISTSOQLTKSAGVGLMGMSHDILAPOTSPILTSVYPTESHF
 YSASATYGSASYSATSCSOLPCGQVQYSVRROQRPBRASBRKSHSESPKQFK
 RNSQMFESJWSNRSRRELGKVSQSSFSGSMEIIVS"

Alignment Scores:

Pred. No.:	0	length:	3059
Score:	472.00	Matches:	663
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	70.98%	Indels:	4
DB:	6	Gaps:	0

US-10-029-345a-109 (1-665) x AX278461 (1-3059)

QY 1 MetAlHieGlueTleGlyThrgInileValThrgInuArgLeuValAlaleuLeuGlu 20
 Db 127 ATGGCCCATGAGATGATTGGACTCAATTTGTTACTGAGAGGTGGGTGCTGCTGGAA 186
 QY 21 SerGlyThrgInuValLeuLeuLleApsSerArgProPheValGluTyraThrSer 40
 Db 187 AGTGAACGGAAAAAGTGTCTAATTGATAGCCGCGCATTTGTGAAATCAATACATCC 246
 QY 41 HieIleuGlunAlaleuLleApsSerArgLeuMetLysArgLeuGluGln 60
 Db 247 CACATTTTGGAAACCATTAATATCACTGCTCAAGCTTAAGAAGCAAGTTGCAACAG 306
 QY 61 AspLysValLeuLleThrgInuLleGlnHieSerAlalyshisLysValAapLleAps 80
 Db 307 GACAAAGCTGTAATTACAGAGCTCATCCAGCATTCAGCGAAACATAAAGTTGACATTGAT 366
 QY 81 CysSerGlnLysValValValTyraApsGlnSerSerGlnApsValAAserLeuSerSer 100
 Db 367 TGCAGTCAGAGAGGTGATTAGTTCAGATCAAAAGCTCCCAAGATGTGCTCTCTCTCA 426
 QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheLysSerValHieLeu 120
 Db 427 GACTGTTTCTCACTGACTTGTGGTAACTGGAGAAAGCTTCAACTGTCTTCACTG 486
 QY 121 LeuAlAGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
 Db 487 CTTCGAGGTGGGTTTGTGCTGCTCTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 546
 QY 141 ThrLeuValProThrCysLieserGlnProCysLeuProValAlaAnuLleGlyProThr 160
 Db 547 ACTCTAGTCCCTACCTGCATTTCTCAGCCTTGCTTACCTGTGCCAAACATTTGGGCCAAC 606
 QY 161 ArgLleLeuProAnuLeuTyLleuGlyCysGlnArgApsValLeuAnuLysGluLeuLle 180
 Db 607 CGAATCTTCCCAATCTTATCTTGGCTGCACAGCATGTCTTCAACAAGAGAGCTGAT- 665
 QY 181 -GlnGlnaenGlyLleGlyTyraValLeuAnuLleAserTyra-ThrCysProLysProAps 200
 Db 666 GCGAGCAAGATGGAGATTGTTATGTTAAATGCGAGAA-TACCTGTCCAAAGCCCTGACT 724
 QY 200 heIleProGlnSerHisPheLeuArgValProValAnuApsSerPheCysGlyLysLleLle 220
 Db 725 TTAATCCCGAGTCTCATTTCTCGGTGCTGCTGATGACACATTTTGTGAGAAAATTT 784
 QY 220 eubProTleuApsLysSerValApsPheLleGlyLysValAlalyshisLysValAps 240
 Db 785 TGGCGGTGTTGCAAAATCAGTATTCATTGAGAAAGCAAAAGCTCCAAATGAGATG 844
 QY 240 alLeuValHisCysLeuAlaGlyLleSerArgSerAlaThrLleAlaLleAlaTyLleLle 260
 Db 845 TTCTAGTGCACTGTTTAGCTGGAGATCTCCGCTCCGACACATCGCTATGCTTACATCA 904
 QY 260 etLysArgMetApsMetSerLeuApsGluAlaTyraArgPheValLysGlyLysArgPro 280
 Db 905 TGAAGAGATGAGCATGTCTTTAGATGACACTTACGATTTGTGAAGAAAAGAACCTTA 964

QY 280 hrLieserProAnuPheAnuPheLeuGlyGlnLeuLeuApsTyrglyLysLysLleLys 300
 Db 965 CATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGACATATGAAAGAAATTAAGA 1024
 QY 300 enGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLleLysLleGlyLysPro 320
 Db 1025 ACCAGACTGAGACATCAGGCGCAAGAGCAAACTCAAGCTGCTGACCTGAGAACCCAA 1084
 QY 320 enGluProAlaProAlaValSerGluGlyGlnLysSerGluThrProLysSerPro 340
 Db 1085 ATGAACCTGTCTCTGTCTGTAAGAGGGTGAAGAGAAAGCAAGAGCCCTCAGTCCAC 1144
 QY 340 roCysAlaApsSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSer 360
 Db 1145 CCGTCCCACTCTGTTACTCTCAGAGGACAGAGCAAAAGCCCGTGCATCCCGCACGG 1204
 QY 360 alProSerValProSerValGlnProSerLeuLeuLleApsSerProLeuValGlnAla 380
 Db 1205 TGGCCAGCGTCCAGCGTGCAGCGCTGCTGTTAGAGAGACAGCCGCTGTACAGGCGC 1264
 QY 380 eusSerGlyLeuHieLysSerAlaApsArgLeuGluApsSerAnuLysLeuLysArgSer 400
 Db 1265 TCAGTGGCTGACCTGTCCGAGACAGGCTGGAAACAGCAATAGCTCAAGCCTTCT 1324
 QY 400 heSerLeuApsLleLysSerValSerTyraSerAlaSerMetAlaAAserLeuHieGly 420
 Db 1325 TCTCTTGATATTCANAATGATTTCTATTTACCGCAGCATGGCAGATCTTACATGCT 1384
 QY 420 heSerSerSerGlnApsAlaLeuGluTyraTyraApsSerThrThrLeuApsGlyThra 440
 Db 1385 TCTCTCATCAGAAAGTCTTTGAAATCACTCAAACTTCACTACTCTGATGGACCA 1444
 QY 440 enLysLeuGlyGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSer 460
 Db 1445 ACAAGCTATGCAAGTCTCTCTGTCAGAACTATTCGAGAGACATCCCGAAACCAAGTC 1504
 QY 460 roApsLysGluGlnAlaSerLleProLysLysLeuGlnThrAlaApsProSerApsSer 480
 Db 1505 CTGATTAAGAGAGAGCAGACATCCCAAGAGCTGCAGACCGCCGCTTCAAGAGCC 1564
 QY 480 InerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 Db 1565 AGAGCAGAGGATTCATTTGCTGTCAGAACAGCAGACAGTGCACCGCCAGAGTCCCTTT 1624
 QY 500 eusSerProLeuHieArgSerGlySerValGluApsAnuTyraHisThrSerPheLeuPhe 520
 Db 1625 TATCTCAGATGCATCGAAGTGGAGCGTGGAGGACAAATTAACACACACACTTCTTTG 1684
 QY 520 LysLeuSerThrSerGlnGlnHieLysLeuThrLysSerAlaGlyLysLeuLysGlyTyra 540
 Db 1685 GCCTTTCCACAGCCAGCAGCACTTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1744
 QY 540 LysSerApsLleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyraPhe 560
 Db 1745 ACTCGATATCTTGGCCCCCAGACTCTTACCTTCTCTGACAGCAGCTGATTTTG 1804
 QY 560 LathrgLysSerHisPheTyraSerAlaSerAlaLleTyrglyLysSerAlaSerTyra 580
 Db 1805 CCACAGAGCTTCAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1864
 QY 580 erAlaTyraSerCysSerGlnLeuProThrCysGlyApsGlnAlaTyraSerValArgArg 600
 Db 1865 CTGCTTACAGCTGACCGCCAGCTGCCACTTGGGAAACCAAGTCTAATTTGTGCGGAGC 1924
 QY 600 rgGlnLysProSerApsArgAlaApsSerArgArgSerTrpHisGluGlnSerProPheG 620
 Db 1925 GGCAGAGCCAGATGACAGAGCTGACTCGCGCGAGCTGCGATGAAGAGAGCCCTTTG 1984
 QY 620 LysGlnPheLysArgArgSerCysGlnMetGluPheGlyGlnSerLleMetSerGln 640
 Db 1985 AAAAGAGTTTAAAGCAGCAAGCTGCAATGGAATTTGAGAGAGCATCATGTCAAGAGA 2044

QY 640 snArGerArGluGluLeuGlyLYeValGlySerGlnSerSerPheSerGlySerMetC 660
|
|
|
Db 2045 ACAGGTCACGGAGAGAGCTGGGAAAGTCGACGCTAGCTTTTCGGGCAGCATGG 2104
QY 660 |u|e|l|e|g|u|a|l|s|e|r 665
|
|
|
Db 2105 AAATCATGTAGGTCTCC 2121
RESULT 9
AX405700 3104 bp DNA linear PAT 14-JUN-2002
LOCUS
DEFINITION Sequence 115 from Patent WO0222660.
ACCESSION AX405700
VERSION AX405700.1 GI:21438839
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
Xue, A.U., Yang, Y., Weinman, T. and Polypeptides
TITLE Novel nucleic acids and polypeptides
JOURNAL Patent: WO 0222660-A 115 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
source Location/Qualifiers
1..3104
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
159..2156
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD34807.1"
/db_xref="GI:21438840"
/db_xref="EMBL:CAD34807"
/translation="MAHEMIGTQIVTERLVALLLESGTEKVLIDSRPFVEYNTSHILE
ATINCSKMKRLOODKYLITELIOHSAKHVDIDSCVVVYOSDVASISDPC
FLTVLLGKLEKSPNSVHILAGAPAFSPFGLCEGSKTVLPTCISOPLDPAVSGT
RLPNILYLCQDVLNKEIMQONGICGYVUNAANTCKPRPIESHFLRPVNDSCCK
ILPMDKSYDFLEKAKASNGCVLVHCLAGISRTATLAVIKRMDMSIDAYRFXKE
KRPTISPNFPLGQLDYERKIKNOGAGSPKSKLLEKPNRPVAPVSGGQKSE
TFLSPCADSATSEAGRPVHPASVPSVQPSLLEDSPLVQALSGIHLSDLEED
SNKRSFSLDIKSVYSASMAASHGFSSSDALEYKPSFTLLDGTNLCOFSPVOS
LSEQTPSPDKPEASIPKKIQTPARSPDSKRIHSVRSSTGTORSLISPLHSGS
VEDNHTSFLFGLSTSQCHLTSAAGLGMKSHDILAPOTSPSITSPSTSPYRTESHP
YASAIYGGASAYSQCLPTCGDQVTSVRROKPPDRADSRKRSWHEESPFEKQPK
RRSCWEPFESIMENSRRELKXGOSQSFSGSMELIEVS"

Db 339 GACAAAGGTTATATACAGAGCTCATCCAGATTCCAGGAAACATTAAGTTACATTGAT 398
QY 81 CysSerGlnLYsValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSer 100
|
|
|
Db 399 TGCACTCAGAAAGTGTAGTTTACATCAAGAGCTCCCAAGATGTGCTCTCTCTTCA 458
QY 101 AspCysPheLeuThrValLeuLeuGlyLYsLeuGlyLYsSerPheAsnSerValHisLeu 120
|
|
|
Db 459 GACTGTTTCTCAGCTGACTTCTGAGTAACTCGAAGAAAGAGCTTCACTGTTTCACTG 518
QY 121 LeuAlaGlyLYsPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLYsSer 140
|
|
|
Db 519 CTTCAGAGTGGGTTTGTGAGTGTCTCTGTTGTTCCCTGGGCTCTGGAAGAAATCC 578
QY 141 ThrLeuValProThrCysAlaLeuSerGlnProCysLeuProValAlaAsnLleGlyProThr 160
|
|
|
Db 579 ACTTAGTCCCTACCTGATTTCTCAGCCTGCTTACCTGTTGCCAACATTTGGGCACAC 638
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyLYsGlnArgAspValLeuAsnLleGlyLeu 180
|
|
|
Db 639 GCAATTTCTCCCAATCTTATCTTGGCTGCCAGCAGATGTCCTCAACAGAGCTGAT- 697
QY 181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProLYsProAsp 200
|
|
|
Db 698 GCAGCAGATGGAGATTGGTTATGTGTTAAATGCCAGCAA-TACCTGTCCAAAGCTGACT 756
QY 200 heileProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlyLYsLle 220
|
|
|
Db 757 TTAATCCCGAGCTCATTTCTGCGCTGCGCTGTGAAGACGCTTTGTGAAAAATTT 816
QY 220 euProTyrLeuAspLYsSerValAspPheIleGlyLysAlaLYsAlaSerAsnGlyCysV 240
|
|
|
Db 817 TGCCGTGTTGGACAAATCAGTAGTTTCATTGAGAAAGCAAAAGCTTCCATGAGATGTG 876
QY 240 a|LeuValHisCysLeuAlaGlyLYsSerArgSerAlaThrIleAlaIleAlaTyrIle 260
|
|
|
Db 877 TTCTAGTCACATGTTTACTGGAGATCTCCGCTCGCACCATCTGATCCGCTACATCA 936
QY 260 e|LYsArgMetAspMetSerLeuAspGluAlaTyrArgPheValLYsGlyLYsAspProT 280
|
|
|
Db 937 TGAAAGAGATGACATGCTTTAGTAGAGCTTACAGATTGTGAAAGAAAAAACCTTA 996
QY 280 hrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLYsLYsIleLYs 300
|
|
|
Db 997 CTATATCTCCCACTTCAATTTTCTGGGCCAATCTTGACATAGAAAGAAATTAAGA 1056
QY 300 snGlnThrGlyAlaSerGlyProLYsSerLYsLeuLYsLeuLeuHisLeuGlyLYsProA 320
|
|
|
Db 1057 ACCAGACTGGAGCATCAGGGCCAAAGACAAACTCAAGCTGCTGCACTGGAGAAAGCCAA 1116
QY 320 snGluProValProAlaValSerGlyLYsGlyGlnLYsSerGlyLthrProLeuSerProP 340
|
|
|
Db 1117 ATGAACCTGTCTCTCTCTCTCTCAGAGGGTGAACAAAAGAGAGGCCCTCACTTCAAC 1176
QY 340 roCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerV 360
|
|
|
Db 1177 CCTGTGCCGACTCTCTACTCTCAGAGCAGCAGCAAGAGCCCTGATCCCGCAGACG 1236
QY 360 a|ProSerValProSerValGlnProSerLeuLeuGlyAspSerProLeuValGlnAla 380
|
|
|
Db 1237 TGCCAGAGGTGCCAGCTGCAGCCGCTGTTAGAGAGACAGCCGCTGGTACAGGCGC 1296
QY 380 euSerGlyLeuHisLeuSerAlaAspArgLeuGlyAspSerAsnLYsLeuLYsAspSerP 400
|
|
|
Db 1297 TCAGTGGCTGACCTGTGCGGAGCAGCTGGAAGACAGCAATAGCTCAAGCCTTCT 1356
QY 400 heSerLeuAspIleLYsSerValSerTyrSerAlaSerMetAlaIleSerLeuHisGlyP 420
|
|
|
Db 1357 TCTCTCTGATATCAATCAAGTTTCAATTTAGCCAGCATGGCAGCACTTACATGGCT 1416
QY 420 heSerSerSerGlyAspAlaLeuGlyTyrTyrLYsProSerThrThrLeuAspGlyThra 440
|
|
|

Db 1417 TCTCTCATCAGAGATGCTTTGGAACTACTACAAACCTTCTCACTACTGATGGACCA 1476
Qy 440 anlyseuCygsInPheserProValGlnGluLeuSerGluGlnThrProGlnThrSerP 460
Db 1477 ACAAGCTATGCGAGCTTCCCTGTTGAGGAAGTATCGAGAGAGACTCCGAAACCAAGTC 1536
Qy 460 roApLySGluGluAlaSerIleProLybLeuGlnThraIaArgProSerAspSerG 480
Db 1537 CTGATTAAGAGAGAACCCAGATCCCAAGAACTGCAAGACCCGAGGCTTCAGACAGCC 1596
Qy 480 InSerIySaArgLeuHISerValArgThrSerSerSerGlyThraIaGlnArgSerLeu 500
Db 1597 AGAGCAAGCGATTGCATTGGCTCAGAACCCAGACAGATGGCAACGCCAAGAGTCCCTTT 1656
Qy 500 euserProLeuHISaArgSerGlySerValGlnuAspAntyRhiSThrSerPheLeuPheG 520
Db 1657 TATCTCACTGCAATCGAAGAGGAGCGTGGAGAGCAATTACACACAGCTTCCCTTTTGG 1716
Qy 520 IyLeuSerThrSerGlnGlnHISerLeuThrIySerSerIaGlyLeuGlyLeuLyGlyTriP 540
Db 1717 GCTTTCCACACACCCAGACAGCACTCAAGAGTCTGCTGGCTGGGCTTTAGGGCTGGC 1776
Qy 540 ISerAspIleLeuAlaProGlnThrSerThrProSerLeuThrIySerSerTriPyrPheA 560
Db 1777 ACTGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGAGAGCTGATTTTG 1836
Qy 560 IahThrGluSerSerIhISpHeTySerAlaSerAlaIleTyGlyGlySerAlaSerTyS 580
Db 1837 CCAACAAGTCTCACTCACTTCTACTCTGCTCAGCCCACTCAAGAGGAGCAGTCACTTACT 1896
Qy 580 eraIaTySerCySeserGlnLeuProThrCyGGLyAspGlnValTySerValArgArgA 600
Db 1897 CTGCTTACAGCTCAGCCAGCTCCCACTTCCGAGACCAAGTCTATTTCTGTCGAGG 1956
Qy 600 rGlnIyAspProSerAspArgAlaAspSerArgIySerTriPyrIaGluGluSerProPheG 620
Db 1957 GGCAGAGCCAGATGACAGAGCTGACTCGCGCGAGCTGGCATGAAGAGGCCCTTTG 2016
Qy 620 IuLyGlnPheLySaArgArgSerCyGlnMetGluPheGlyGluSerIleMetSerGluA 640
Db 2017 AAAAGCAGTTTAAACCCAGAGAGCTGCCAAATGCAATTTTGAGAGACATCATGTCAAGA 2076
Qy 640 snArgSerArgGluGluLeuGlyLySaValGlySerGlnSerSerPheSerGlySerMetG 660
Db 2077 ACGAGTCAAGGAGAGCTGGGAGAGTGGGAGCTAGCTTACCTTTTCGGGACGATGG 2136
Qy 660 IuIleIleGluValSer 665
Db 2137 AATCATTTAGGCTCC 2153

RESULT 10
AX441229 3332 bp DNA linear PAT 28-JUN-2002
LOCUS AX441229
DEFINITION Sequence 20 from Patent WO0226997.
ACCESSION AX441229
VERSION AX441229.1 GI:21665771
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS LucHe,R.M. and Wei,B.
TITLE Dep-16 dual-specificity phosphatase
JOURNAL Patent: WO 0226997-A 20 04-APR-2002;
Ceptlyr, Inc. (US)
FEATURES
source location/Qualifiers
1. .3332
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 3332
Score: 472.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.98% Indels: 0
DB: 6 Gaps: 0
US-10-029-345a-109 (1-665) x AX441229 (1-3332)

Qy 194 ThrCySProLyAspPheIleProGluSerHISpHeLeuArgValProValAsnAsp 213
Db 977 ACTGTCGAAGCGCTGACTTATCCCGAGTCTCATTTCTCGTGGTGCCTGGAATGAC 1036
Qy 214 SerPheCyGlnGluIleLeuProTriPyrLeuAspLybSerValAspPheIleGlyLySaIa 233
Db 1037 ACCTTTTGAGAAAAATTTTGGCGTGGAGCAATAGTGAATTTCAATTGGAAGA 1096
Qy 234 IySaIaSerAspGlyCySValLeuValHISCySLeuAlaGlyIleSerArgSerAlaTrn 253
Db 1097 AAGCTCCAAATGATGATGTCTTAGTGCACTGTTTACTGGGATCTCCGCTCCGACAC 1156
Qy 254 IleaIaIleAlaTyIleMetLySaArgMetAspMetSerLeuAspGluAlaTyArgPhe 273
Db 1157 ATCGTATCGCTTACATCATGAAGAGATGACATGTCTTTAGATGAACCTTACAGATT 1216
Qy 274 ValLybGluLybArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAsp 293
Db 1217 GTGAAGAAAAAAGAACTCTATCTATCTCAAACTTCATTTCTGGGCAACTCTCGAG 1276
Qy 294 TyGlyLybLybIleLybAsnGlnThrGlyAlaSerGlyProLybSerLybLeuLybLeu 313
Db 1277 TATGAGAAAGAAATTAGAACCAAGCTGAGATTCAGGGCCAAAGCAAACTCAAGCTG 1336
Qy 314 LeuHISLeuGluLybProAsnGluProValProAlaValSerGluGlyGlyGlyLybSer 333
Db 1337 CTGACCTGGAAGAGCAATGAACTGTCCCTGCTGTCAAGAGGTGGAAGAAAGC 1396
Qy 334 GluThrProLeuSerProProCybAlaAspSerAlaThrSerGluAlaIleGlyGlnArg 353
Db 1397 GAGACCCCTCAAGTCACTCACCTGTGCGCACTGTCACTCACTCAAGAGCAGACAGCAAA 1456
Qy 354 ProValHISProAlaSerValProSerValProSerValGlnProSerLeuLeuGluAsp 373
Db 1457 CCGTGATCCCGGACAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAG 1516
Qy 374 SerProLeuValGlnAlaLeuSerGlyLeuHISLeuSerAlaAspArgLeuGluAspSer 393
Db 1517 AGCCGCTGTGTAAGGCGCTCAAGGCTGCACTGTCCGACAGACAGGCTGGAAGACAGC 1576
Qy 394 AsnLybLeuLybArgSerPheSerLeuAspIleLybSerValSerTySerAlaSerMet 413
Db 1577 AATAAGCTCAACAGTTCCTTCTCTGATTAACAATGATTTCATATTCAGCCACACAG 1636
Qy 414 AlaIaSerLeuHISGlyPheSerSerSerGlyuAspAlaLeuGlnuTyRlybProSer 433
Db 1637 GAGCATCTTACATAGCTTCTCTCATAGAAAGCTTTGGAAATCAACAACTTCC 1696
Qy 434 ThrThrLeuAspGlyThraAsnLybLeuCyGlnPheSerProValGlnGluLeuSerGlu 453
Db 1697 ACTACTCTGATGGAGCAACCAAGCATATGCCAGTTCTCCCTGTTCAAGAACTATGGAG 1756
Qy 454 GlnThrProGluThrSerProAspLybGluGluAlaSerIleProLybLybLeuGlnThr 473
Db 1757 CAGACTCCCGAAACCAAGCTGTATAGAGGAGGAGCAGCATCCCAAGAGCTGCGAGACC 1816
Qy 474 AlaArgProSerAspSerGlnSerLybArgLeuHISerValArgThrSerSerSerGly 493
Db 1817 GCGAGGCTTTCAGACGACGAGCAAGCGATTGCACTCGGTCAAGAACGACGACAGTGGC 1876
Qy 494 ThrAlaGlnArgSerLeuLeuSerProLeuHISaArgSerGlySerValGluAspAntyR 513
Db 1877 ACCGCCCAAGGTCCTTTTATCTCAGCTGATCAAGAGTGGAGCGTGAAGACAAATTAC 1936

QY 514 HisThrSerPheIeuPheGlyIeuSerThrSerGlnGlnHisIeuThrIysSerAlaGly 533
Db 1937 CACACCAAGCTTCCTTTTGGCCCTTTCCACAGCAGACACCTCAGAACTGCTGGC 1996
QY 534 LeuGlyIeuIysGlyIleThrIleSerAspIleLeuAlaProGlnThrSerThrProSerLeu 553
Db 1997 CTGGGCTTAAAGGCTGGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTG 2056
QY 554 ThrSerSerTrpIlePheAlaThrGluSerSerHisPheTyIserAlaSerAlaIleTy 573
Db 2057 ACCAGAGCTGTATTTTGGCAGAGACTCTCACACTTACTCTGCTCAGCCATCTAG 2116
QY 574 GlnGlySerIleSerIleTyIserAlaTyIserCysSerGlnIeuProThrCysGlyAspGln 593
Db 2117 GGGAGGAGTCCAGTTCCTGCTCAGCTGACAGCTGAGCTGCCCTGGCAGACCA 2176
QY 594 ValTyIserValArgArgGlnIysProSerAspArgAlaAspSerArgArgSerTrp 613
Db 2177 GTCTATTCGTGGCGAGCGCGGAGAACCAAGTGAAGAGCTGACTCGCGCGAGCTGG 2236
QY 614 HisGlnGluSerProPheGlnIysGlnPheIysArgArgSerCysGlnMetGlnPheGly 633
Db 2237 CATGAAGAGAGCCCTTTGAAAAGCAGTTTAAACGACGAAGCTGCCAAATGGATTGGGA 2296
QY 634 GluSerIleMetSerGlnAsnArgSerArgGlnGluLeuGlyIleValGlySerGlnSer 653
Db 2297 GAGAGCATCATGTCAAGAACAGCTCACGGAGAGAGCTGGGAGAACTGGGCGAGTCTC 2356
QY 654 SerPheSerGlySerMetGlnIleIleGlnValSer 665
Db 2357 AGCTTTTCGGGCGAGCATGGAAATCATTTGAGTCTCC 2392

RESULT 11
AX441210
LOCUS AX441210 3496 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 1 from Patent WO0226997.
ACCESSION AX441210
VERSION AX441210.1 GI:21665766
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
1 Luche, R.M. and Wei, B.
Dsp-16 dual-specificity phosphatase
Patent: WO 0226997-A 1 04-APR-2002;
Cepcyr, Inc. (US)

FEATURES
source 1..3496
location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3496
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 6 Gaps: 0

US-10-029-345A-109 (1-665) x AX441210 (1-3496)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValIleThrGluArgLeuValAlaIeuLeuGln 20
Db 562 ATGGCCCATGAGATGATTGGAACTCAAAATGTTACTGAGAGGTTGGTGGCTCTGTGGA 621
QY 21 SerGlyThrGlnIysValIeuLeuIleAspSerArgProPheValGlnTyIAsnThrSer 40
Db 622 AGTGAACGGAAAAAGTGTCTAATGTATGATGACCGGCATTTGTGAATATACATATCCTC 681

QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerIysIleMetIysArgArgLeuGln 60
Db 682 CACATTTTGAAGCCATTAAATATCACTGCTCCAGCTTATGAAAGGTTGCAACAG 741
QY 61 AspIysValIleIleThrGlnIeuIleGlnHisSerAlaIysHisIysValAspIleAsp 80
Db 742 GACAAAGTTTAAATACAGAGCTCATCCAGCATTCAGGAAACATTAAGGTTGACATTAAT 801
QY 81 CysSerGlnIysValIleValTyIAspGlnSerSerGlnAspValAlaSerIeuSerSer 100
Db 802 TGCAGTCAAGAGGTTGATGTTACATCAAGAGCTCCCAAGATGTTGCCCTCTCTCTTCA 861
QY 101 AspCysPheIeuThrValIeuLeuGlnIysIleuGlnIysSerPheAsnSerValHisIeu 120
Db 862 GACTGTTTTCACAGTCACTTCTGGGTAACTGGAGAAAGACTTCACTGTTTCACTGG 921
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyIleuGlnGlyIysSer 140
Db 922 CTTCAGAGTGGGTTTGTGAGTTCCTGTTGTTTCCCTGGCCTCTGTGAAGAAATTC 981
QY 141 ThrIeuValProThrCysIleSerGlnProCysIeuProValAlaAsnIleGlyProThr 160
Db 982 ACTGATGCTCCTACCTGATTTCTCAGCTTGCTTACCTGTTGCAACATTTGGGCAAC 1041
QY 161 ArgIleIeuProAsnIeuTyIleuGlyCysGlnArgAspValIeuAsnIysGlnIeuIle 180
Db 1042 CGAATTCCTCCCACTTTATCTTGGCTGACGAGAGATGTCCTCAACAGAGCTGAT- 1100
QY 181 -GlnIleAsnGlyIleGlyTyIValIleuAsnAlaSerTyI-ThCysProIysProAsp 200
Db 1101 GCAGCAGATAGGATGTTGTTATGCTTAAATGCCAGCA-TACCTGTCCAAAGCTTGACT 1159
QY 200 IleIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlnIysIle 220
Db 1160 TTATCCCGAGTCTCATTTCTGCGTGGCTGTGAATGACAGCTTTTGTGAAGAAATTT 1219
QY 220 EupProTPrIeuAspIysSerValAspPheIleGlnIleValIleSerAsnGlyCysV 240
Db 1220 TGCCTGTGTGAAATAATCACTAATATTCATTTGAAAGCAAAAGCTCCCATATGATGTG 1279
QY 240 AlIeuValHisCysIleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyIle 260
Db 1280 TTCTAGTCACTGTTTACCTGGATCTCCGCTCCGACCATGCTATGCGCTTACATCA 1339
QY 260 eIysArgMetAspMetSerIeuAspGlnAlaTyIArgPheValIysGlnIysArgProT 280
Db 1340 TGAAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAGAAAGACTTA 1399
QY 280 hrIleSerProAsnPheAsnPheIeuGlyIleuLeuAspTyIArgIleIysIleIysA 300
Db 1400 CTATATCTCCAAACTTCAATTTCTTGGGCCCAACTCTGAGCTATGAAGAAAGATTAGA 1459
QY 300 snGlnThrGlyAlaSerGlyProIysSerIysIeuIysIleuHisIleuGlnIysProA 320
Db 1460 ACCAGACTGAGATTCAGGGCCAAAGCAAACTCAAGCTGCTGACCTGGAAGAGCCAA 1519
QY 320 snGlnProValProAlaValSerGlnIysGlnIysSerGlnIysThrProIeuSerProp 340
Db 1520 ATGAAGCTGTCCTGCTGCTCAGAGGGGTGACGAAAGACGAGAACCCCTCAGTCAC 1579
QY 340 rocYsaAlaAspSerAlaThrSerGlnAlaIleGlyIleuArgProValHisProAlaSerV 360
Db 1580 CTGTGCGGACTGCTGACTCAGAGGACAGAGCAAGAGCCCTGGTCACTCCGCCAGCG 1639
QY 360 alProSerValProSerValGlnProSerIeuLeuGlnIysAspSerProIeuValGlnAla 380
Db 1640 TGCCACAGCTGCCCAAGCTGACGCTGCTGTTAAGAGCAAGCCGCTGTACAGGCGC 1699
QY 380 eusSerGlyIleuHisIeuSerAlaAspArgIleuGlnIysSerAsnIysIleuIysArgSerP 400
Db 1700 TCAGTGGGTGACACTGTCCGACAGACAGGCTGGAAGACACAAATTAAGCTCAACGTTCT 1759

QY 400 hseSerleuapbilleysSerValSerTySerSerAlaSerMetAlaAlaSerLeuHleuGlyP 420
 DB 1760 TCTCTGTGATATCAATCAATGTTTATATTCAGCCAGCATGAGAGATCTTATCAATGAGCT 1819
 QY 420 hseSerSerSerGluAapAlaLeuGluTyTyTySerProSerThrThrLeuAapGlyThra 440
 DB 1820 TCTCCCTCATCAGAAAGATGCTTGGATACACAAACCTTCCACTGATCTGATGGAGACA 1879
 QY 440 gnllyleuGlySerGlnPheSerProValGlnGluLeuSerGluGlnThrProGlnThrSerP 460
 DB 1880 ACAAGCTAATGCCAGTCTCCCTCCCTCAGAAATATCGAGAGAGAGAGAGAGAGAGAGAG 1939
 QY 460 roaAaplysgluGluAlaSerIleProLylyleuGlnThraAapProSerAapSerG 480
 DB 1940 CTGATAG 1999
 QY 480 lnseryaArgLeuHleuHleuHleuHleuHleuHleuHleuHleuHleuHleuHleuHleu 500
 DB 2000 AGAGCAAG 2059
 QY 500 euseSerProleuHleuHleuHleuHleuHleuHleuHleuHleuHleuHleuHleuHleu 520
 DB 2060 TATCTCCTCAGTCAATCAAG 2119
 QY 520 lyleuSerThrSerGlnGlnHleuHleuHleuHleuHleuHleuHleuHleuHleuHleu 540
 DB 2120 GCTTTTCCACACAG 2179
 QY 540 laseSerAplleleuAlaProGlnThrSerThrProSerLeuThrSerSerTyTyTyPhea 560
 DB 2180 ACTCGATATCTTGGCCCCAG 2239
 QY 560 lathrGluSerSerHleuHleuHleuHleuHleuHleuHleuHleuHleuHleuHleuHleu 580
 DB 2240 CCACAG 2299
 QY 580 eAlaTySerSerTySerSerGlnLeuProThrCyGlyAapGlnValTySerValAapAra 600
 DB 2300 CTGCTTACAGCTGAG 2359
 QY 600 rsgGlnlyAapProSerAapAraAlaAapSerAraAapSerTyTyTyTyTyTyTyTyTyTy 620
 DB 2360 GGCAG 2419
 QY 620 lulyeGlnPheTy 640
 DB 2420 AAAAG 2479
 QY 640 sAapSerAraAapGlnGluLeuGlyValGlySerGlnSerSerPheSerGlySerMetG 660
 DB 2480 ACGAGTCAAG 2539
 QY 660 lulleleGluValSer 665
 DB 2540 AATCATTTAGAGTCTCC 2556
 RESULT 12
 AB052156 3521 bp mRNA linear PRI 18-OCT-2001
 LOCUS Homo sapiens MKP-7
 DEFINITION Homo sapiens MKP-7 mRNA for MAPK phosphatase-7, complete cds.
 ACCESSION AB052156
 VERSION AB052156.1 GI:13548676
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 Masuda, K., Shima, H., Watanabe, M. and Kikuchi, K.
 TITLE MKP-7, a novel mitogen-activated protein kinase phosphatase,
 functions as a shuttle protein
 JOURNAL J. Biol. Chem. 276 (42), 39002-39011 (2001)
 MEDLINE 21486429

PUBMED 11489891
 2 (bases 1 to 3521)
 REFERENCE
 AUTHORS Masuda, K., Shima, H. and Kikuchi, K.
 TITLE Direct Submission
 JOURNAL Submitted (05-DEC-2000) Kouhei Masuda, Institute for Genetic
 Medicine, Division of Biochemical Oncology and Immunology, Kita-ku
 Kitas-5-jou nishi1-kyoume, Sapporo, Hokkaido 060-0815, Japan
 (E-mail:kou@imm.hokudai.ac.jp, Tel:81-11-706-5536,
 Fax:81-11-707-6839)
 FEATURES
 source
 1. 3521
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="12"
 /cell_line="Jurkat"
 1. 3521
 /gene="MKP-7"
 564. 2561
 /gene="MKP-7"
 /codon_start=1
 /product="MAPK phosphatase-7"
 /protein_id="BAB40814.1"
 /db_xref="GI:13548677"
 /translation="MAHEMIGQIVTERVALLBSGTEKVLIDSRPFVEVNTSHILE
 AINISKLMKMRRLQODKVLITLLOHSAKHVDIDGQKVVVYVDSQDVASLSDC
 FLTVILGLBKSFNVLHLAGFAPSRCPFLCEKSTIVPTCLISOPCLPANTIGPT
 RLPLNLYGCGQDVNLKELMOQNGIGVYLNASNTCPKDFIPESHPLRPVNDSCFEK
 ILPLWLDKVDPIEKAKASNGCVLHCLNGISKSATIALAYIKRMKDMSDAEYRFEKE
 KRPTISPNPFLGQLDYEKIKNOTGASGPKSKLILKRNPNPVPVAVSGQKSE
 TPLSPCADSATSEAAAGORPVHPVAPSVYOPSLLEDSPLVQALSGHLADRLPE
 SNKLRSPLDIDKSYSVASMAASLHGSSSEDALEYKPPSTTLGNTLCCPSPVQ
 LSEQPERTSPDKEASTIKKLOTARPSPQSQRLLSVRTSSSGTRQSRLLSLHNSGS
 VEDNHTSFLFGLSTSQHLTASLGLKMHSDILAPQTSPLSLTSVYFATESHF
 YSAASLYGSAVSAYSCSLPTCDQOVYSVARRKQPSDRADSRMSWRESPFEKQFK
 RSSQMEGEGESIMSENREBELKVGSSSFEFGSMELIEVS"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3521
 Score: 472.00 Matches: 663
 Percent Similarity: 99.40% Conservative: 0
 Best Local Similarity: 99.40% Mismatches: 2
 Query Match: 70.98% Indels: 4
 DB: 9 Gaps: 0
 US-10-029-345A-109 (1-665) x AB052156 (1-3521)
 QY 1 MetAlaHleuGlnMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 DB 564 ATGGCCCATGAGATGATGGAACTCAAAATTTGATCTGAGAGGTTGGCTCTGCTGGAA 623
 QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 DB 624 AGTGAAGCGAAGAAAGTGTCTAATTTGATACCCGCCATTTGTGGAATACATCATC 683
 QY 41 HleleleuGluAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 60
 DB 684 CACATTTTGAAGCATTATATCAATGCTCCAGCTTATGAGGAGGAGGAGGAGGAGGAGGAG 743
 QY 61 AaplyValleuIleThrGluLeuIleGlnHleuHleuHleuHleuHleuHleuHleuHleu 80
 DB 744 GACAAAGTITTAATTCAGAGCTCATTCAGATTGAGCCAAACATTAAGGTGATTCATTGAT 803
 QY 81 CySerGlnlyValValValValValTyraAapGlnSerSerGlnAapValAlaSerLeuSerSer 100
 DB 804 TCGAGTCAAGAGAGTGTAGTTACATCAAACTCCCAAGAGAGTGGCTCTCTCTTCA 863
 QY 101 AapCyPheLeuThrValleuLeuGlylylyleuGluTyTyTyTyTyTyTyTyTyTyTyTy 120
 DB 864 GACTGTTTTCTCAGTACTTCTGAGTAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 923
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerAraGlyPheProGlyLeuCyGlyGluTyTySer 140

Db 924 CTGGACAGTGGATTGCTGAGTTCTCTCGTTGTTCCCTGCTCTGTGAAGAAATTC 983
 Qy 141 ThrluValProthrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProthr 160
 Db 984 ACTCTAGTCCCTCAGCTGATCTTCTCAGCCTTGTCTTACCTGCTTCCAAATGGGCCAAC 1043
 Qy 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGlyLeu 180
 Db 1044 CGAATCTTCCCAATCTTATCTTGGCTGCCAGAGATGCTCTCAACAAAGAGCTGAT- 1102
 Qy 181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProLysProAsp 200
 Db 1103 GCAGCAGAATGGGATTGTTATGTTAAATGCCAGAA-TACTCTTCCAAAGCCTGACT 1161
 Qy 200 helLeuProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIle 220
 Db 1162 TTATCCCCAGTCTCATTTCTCTGCTGCTGTGATGACAGCTTTTGTGAGAAATTT 1221
 Qy 220 euProTTrpLeuAspLysSerValAspPheIleGlnLysAlaLysAlaSerAsnGlyCys 240
 Db 1222 TGCCGTGGTTGGCAAAATCGTAGATTTCATAGAAAGCAAAAGCTCCAAATGAGTGG 1281
 Qy 240 alLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 260
 Db 1282 TTCTAGTCACTGTTAGCTGGGATCTCCCGCTCCGCCAACATCGCTATCGCTACATCA 1341
 Qy 260 euLysArgMetAspMetSerLeuAspGlnAlaTyrArgPheValLysGlnLysValGly 280
 Db 1342 TGAAGAGATGACATGCTTTAGATGAAAGCTTACGATTGTGAAAGAAAGAAAGCTTA 1401
 Qy 280 hrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnLysIleLys 300
 Db 1402 CTATATCTCCAAATCTCAATTTTCTGGGCCAACTCTGAGCATATGAGAAAGATTAA 1461
 Qy 300 snGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlnLysPro 320
 Db 1462 ACCAGACTGGAGATCATGGGCCCAAGAGCAAACTCAAGCTGCTGACCTGGAGAAAGCCA 1521
 Qy 320 snGlnProValProAlaValSerGlnGlyGlnLysSerGlnThrProLeuSerPro 340
 Db 1522 ATGAACTTCTCCCTGCTGCTCAGAGGGTGAAGAGAAAGCCAGAGCGCTCACTCAC 1581
 Qy 340 roCysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSer 360
 Db 1582 CCGTGGCCGACTCTGTACTCAGAGGCGAGAGCAAAAGCCCTGATCCCGCCAGCG 1641
 Qy 360 alProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAla 380
 Db 1642 TGCCCAAGCTGGCCAGCGTGCAGCGCTGCTGTAGAGAGACGCCGCTGTACAGCGCG 1701
 Qy 380 euSerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSer 400
 Db 1702 TAGAGTGGCTGCACTGTCGAGAGAGGCTGGAAGACGCAATAGCTCAAGCCTTCT 1761
 Qy 400 heSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGly 420
 Db 1762 TCTCTCTGGATATCAATCAGTTTCAATATTCAGCCAGCATGSCATCTTACATGGCT 1821
 Qy 420 heSerSerSerGlnAspAlaLeuGlnTyrTyrLysProSerThrThrLeuAspGlyThr 440
 Db 1822 TCTCTCTCATCAAAATGCTTGTGAATATCAAAACCTTCCACTCTCGATGGAGACA 1881
 Qy 440 snLysLeuCysGlnPheSerProValGlnLeuSerGlnGlnThrProGlnThrSer 460
 Db 1882 ACAAGCTATGCCAGTTCCTCTGTTCCAGAACTATCGAGAGAGCTCCCAAAACCAATC 1941
 Qy 460 roAspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSer 480
 Db 1942 CTGATTAAGAGAGAACCGACGATCCCAAGAGCTGCAACCGCCAGGCTTTCAGACGCC 2001
 Qy 480 InSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500

Db 2002 AGAGCAAGCATTCGATTCGCTCAGAACGACGATGGCACCCGCCCAAGATTCCTTT 2061
 Qy 500 euSerProLeuHisArgSerGlySerValGlnAspAsnTyrHisPhePheLeuPhe 520
 Db 2062 TATCTCCACTGCATGAGAGTGGAGAGCTGGAGAGCAATATCACACCGACTTCTTTCG 2121
 Qy 520 lLysLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysLeuGly 540
 Db 2122 GCTTTTCCACGACGACGACGACCTTCAGAAAGTGTGCTGGCTGGCTTAAAGGCTGGC 2181
 Qy 540 lSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpIlePhe 560
 Db 2182 ACTCGATATCTGGCCCCCAGACCTTACCTTCCCTACACGACGCTGGATTGTTG 2241
 Qy 560 lArgGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyr 580
 Db 2242 CCACAGAGTCTCCACACTTCTACTCTGCTGACGACATCTACGAGGAGGAGTCCAGTTACT 2301
 Qy 580 erAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
 Db 2302 CTGCTTACAGCTGACGACGCTGCCACTTGCAGAGACCAAGTATTTCTGTGGCGGC 2361
 Qy 600 rGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGlnGlnSerProPhe 620
 Db 2362 GCGAGAACCAAGTACAGAGACTGATCGCGCGAGCTGGCATGAGAGAGCCCTTTG 2421
 Qy 620 lLysGlnPheLysArgArgSerCysGlnMetGlnPheGlyLysSerIleMetSerGln 640
 Db 2422 AAAAGCATTTTAAACGCAAGAGCTGCCAAATGAAATTTGGAGAGACATCATGTCAAGA 2481
 Qy 640 snArgSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMet 660
 Db 2482 ACAAGTCAAGGAGAGAGCTGGGAGAAAGTGGGACGTAGCTTACTTTTGGGACAGATGG 2541
 Qy 660 lLleIleGlnValSer 665
 Db 2542 AAATCATTAGAGTCTCC 2558
 Db 2558
 RESULT 13
 AX260340
 LOCUS 3544 bp DNA linear PAT 26-OCT-2001
 DEFINITION Sequence 1 from Patent WO0173059.
 ACCESSION AX260340
 VERSION AX260340.1 GI:16509303
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1
 Meyers, R.A.
 38692 and 21117: dual specificity phosphatase molecules and uses
 thereof
 Patent: WO 0173059-A 1 04-OCT-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source
 location/Qualifiers
 1..3544
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 589..2586
 /note="unamed protein product"
 /codon_start=1
 /protein_id="CAD10218.1"
 /db_xref="GI:16509304"
 /db_xref="REMTREMBL:CAD10218"
 /translation="MAHEMIGQIVTERIVALLSSTGEKVLIDSRPFVENVTSILIE
 AININSKIMKRRLOODKVLITELIOHSAKHVDIDCSQKVVYVYDSSQDVASISDC
 RILPLVLGKLEKSPNSYHLAAGFAESRFRPGICEKSKSLVPRCISQPLPIANISPT
 RILPNIYLGGQVRLNKEMLQOQVIGTIVLNASATCKPPIPSHPLRVVNDSPCEK
 ILFWIDKSDVFLERKASNGCYLVHGLAISRAITLAIVYIKRMDMSIDEAYRVFYE
 KRPTISPNFNFDQILDYERKIKNOTGASGSPSKLKLHLERNEPVPVAVSGGQKSE
 TPLSPCADSATSSEAGRPVHPASVPSVQPSILBDSPIVGLSLGLHSADRLIED

ORIGIN

SNLKRKPSFLDIKSVSYASAMASLHGFSSSEDALEYKSTFLDITDINKLCOFSPVOR
LSEQTPETSPDKEBASIPKLTOTARPSDSOKRHSVRFTSSSGTORSLISPLHSGS
VEDNHTTSLRGLSTSOCHLTYSAGLGMKSHDIAPOPTSPSLTSSRYFTSHSHF
YASASATYGSASISATYSCQPLTCGDQVTSVRROKPSRROASRHSWBESEPEROKPK
RSCQMEFESIMSENRSERELKGVSSQSSFSGSMELIIVS"

Alignment Scores:

Pred. No.:	0	Length:	3544
Score:	472.00	Matches:	663
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	70.98%	Indels:	4
DB:	6	Gaps:	0

US-10-029-345A-109 (1-665) x AX260340 (1-3544)

QY 1 MetAlaHieGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB ATGGCCCATGAGATGATGGAAGTCAATTTGTTACTGAGAGATGGTGGCTCTGCTGGAA 648
QY 21 SerGlyThrGluLeuValLeuLeuIleAspSerArgProPheValGluTyrAnthrSer 40
DB AGTGAACCGGMAAAGTGGCTGATTAATGATGACCGGCAATTTGTGGAATACATACATCC 708
QY 41 HisIleLeuGluAlaIleAniIleAsnCySerLeuLeuMetIleArgProLeuGlnGln 60
DB CACATTTTGAAGCCATTAATATCAACCTGCTCCAGCTTAATGAGGAAAGTTGGCAACG 768
QY 61 AspLeuValLeuIleThrGluLeuIleGlnHisSerAlaIleHisLeuValAspIleAsp 80
DB GACAAAGTGTATTAATACAGAGCTCATCCAGCATTCAGCAACATTAAGTTGACATTTGAT 828
QY 769 GACAAAGTGTATTAATACAGAGCTCATCCAGCATTCAGCAACATTAAGTTGACATTTGAT 828
DB 81 CySerGlnLeuValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB TGCAGTCAGAGAGTTGATTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCTCA 888
QY 829 TGCAGTCAGAGAGTTGATTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCTCA 888
DB 101 AspCyPheLeuThrValLeuLeuGlyIleValLeuGluLeuSerPheAnthrValHisLeu 120
DB 889 GACTGTTTCTCACTGACTCTGGGTAAACGTGAGAAAGCTTCAACTCTGTTACCTG 948
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCyPheProGlyLeuCyGluGlyIleSer 140
DB 949 CTTCGAGGTGGGTCTCTGATCTCTCTGTTTCCCTGGCTCTGTGAAGAAATTC 1008
QY 141 ThrLeuValProThrCySerIleSerGlnProCyLeuProValAlaAniIleGlyProThr 160
DB 1009 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTCACCAATTGGGCAACC 1068
QY 161 ArgIleLeuProAnthrLeuGlyCyGlnArgAspValLeuAniIleGlyLeuIle 180
DB 1069 CGAATCTTCCCAATCTTATATCTTGGCTGCACGAGATGCTCTCAACAGAGCTGAT- 1127
QY 181 -GlnGlnAniGlyIleGlyTyrValLeuAniIleSerTyr-ThrCySProLysProAsp 200
DB 1128 GCAAGCAATGGATGGTATGTATGTAAATGCCAGCA-TACCTGTCCAAAGCCGAT 1186
QY 200 HeIleProGlnSerHisPheLeuArgValProValAniAspSerPheCyGluValIle 220
DB 1187 TATATCCCAAGTCTCAATTTCTGCTGCTGTGATGATGACAGCTTTGTGAGAAATTT 1246
QY 220 eulProThrLeuAspLysSerValAspPheIleGluValAlaValAspAniIleCyav 240
DB 1247 TGCCTGGTGGTGAACATAGATATTCATTGAGAAAGCAAAAGCTCCCAAGATGAG 1306
QY 240 alLeuValHisCySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleM 260
DB 1307 TTCTAGTCACTTTTACGTGGGATCTCCGCTCCGCCACATCGCTTACGCTTACATCA 1366
QY 260 eulValArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluValArgPro 280
DB 1367 TGAAGAGATGACATGTCTTATGATGAGCTTACAGATTTGTGAGAAAGAAAGACCTTA 1426

QY 280 hrlIleSerProAnthrPheAnthrPheLeuGlyGlnIleLeuAniIleValIleVal 300
DB 1427 CTAATCTTCCAAACTTCAATTTTCTGGGCAACTCTGACATATGAGAAAGATTAAGA 1486
QY 300 eulIleThrValIleSerGlyProLysSerValLeuValLeuLeuIleValLeuGluValPro 320
DB 1487 ACCAGACTGAGACATCAAGGCAAGAGCAAACTCAAGCTGTGACCTGGAGAAAGCCAA 1546
QY 320 eulProValProAlaValSerGluGlyGlyGlnIleValSerGlnThrProLeuSerPro 340
DB 1547 ATGAACCTGTCCCTCTCTGCTCAGAGGGTGAACAAAAAGCAAGAGCCCTCAGTCCAC 1606
QY 340 roCyAlaAspSerAlaThrSerGluAlaAlaGlyIleValArgProValHisProAlaSerV 360
DB 1607 CTTGTGCCACTCTCTCTACTTACCTTACAGGCGAGAGCAAGCAAGCCCTGTGATCCCGCAGG 1666
QY 360 alProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAla 380
DB 1667 TGCAGAGGTGCCAGCGTGCAGCCCTGCTGTAGAGACAGCCGCTGTACAGAGCCG 1726
QY 380 eulSerGlyLeuHisIleLeuSerAlaAspArgLeuGluAspSerAniIleValArgSerP 400
DB 1727 TCAGTGGGCTGCACCTGTCCCGACAGAGCTGGAAGACAGCAATTAAGCTCAAGCTTCT 1786
QY 400 HeSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaIleSerLeuHisGlyP 420
DB 1787 TCTCTGTGATATCAATATGATTTATATTCAGCCAGATGACAGATCTTACATGCT 1846
QY 420 HeSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThr 440
DB 1847 TCTCTCATCAAGAGTGTGTTGAAATCTACAAACCTTCGACTCTGTGATGGAGCA 1906
QY 440 eulValLeuCyGlnPheSerProValGlnIleLeuSerGluGlnThrProGluThrSerP 460
DB 1907 ACNACGTATGCGAGTCTCCCTGTTCAGGAATCTGAGAGAGACTCCCGAAACAGAGTC 1966
QY 460 roAspLeuGluGlnIleAspIleProLysValLeuGlnThrAlaArgProAspAspSerG 480
DB 1967 CTGATTAAGAGAGAGAGAGAGATCCCGAGAGCTGACAGCCCGAGCTTTACAGAGCC 2026
QY 480 InSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
DB 2027 AGAGCAGAGATGATTCATTCGTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2086
QY 500 eulSerProLeuHisArgSerGlySerValGluAspAniIleThrSerPheLeuPheG 520
DB 2087 TATCTTCACTGATGAAAGTGGAGAGTGGAGAGCAATTAACACACACACTTCTTTTG 2146
QY 520 ILeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrP 540
DB 2147 GCTTTTCCAGCCAG 2206
QY 540 IAspAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrTyrPhe 560
DB 2207 ACTCGGATATCTTGGCCCCCAGAGCTTCAACCTTCCCTGACAGAGAGAGAGAGAG 2266
QY 560 IATThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrS 580
DB 2267 CCACAGAGAGCTTCAACATTTCTGCTGCTGACCATCAACAGAGAGAGAGAGAGAG 2326
QY 580 eulAlaTyrSerCySerGlnLeuProThrCyGlyAspGlnValTyrSerValArgArg 600
DB 2327 CTGCTTACAGCTGAG 2386
QY 600 rglIleLysProSerAspArgAlaAspSerArgArgSerTyrHisGluGluIleProPheG 620
DB 2387 GGCAGAGAGCAAGTACAGAGAGTGAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2446
QY 620 IuLysGlnPheLysArgArgSerCyGlnMetGluPheGlyGluSerIleMetSerGlu 640
DB 2447 AAAAGCAGATTTAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2506
QY 640 anArgSerArgIleGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetG 660

|||||
Db 2507 ACAGGTACCGGAAGACTGGGAAAGTGGCACTGACTTTCGGGACGATGG 2566
Qy 660 lullelleglualSer 665
Db 2567 AAATCATTGAGGTCTCC 2583
RESULT 14
AF506796 3566 bp mRNA linear PRI 30-OCT-2003
LOCUS AF506796
DEFINITION Homo sapiens dual specificity phosphatase 16 (DUSP16) mRNA,
complete cds.
ACCESSION AF506796
VERSION AF506796.1 GI:25573087
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 3566)
TITLE MAPK phosphatase DUSP16/MKP-7, a candidate tumor suppressor for
chromosome region 12p12-13, reduces BCR-ABL-induced transformation
JOURNAL Oncogene 22 (49), 7728-7736 (2003)
PubMed 14586399
REFERENCE 2 (bases 1 to 3566)
AUTHORS Hoonraert, I., Marynen, P., Goris, J., Sciote, R. and Baens, M.
TITLE Submitted (26-APR-2002) Department for Human Genetics-Flanders
Internuversity Institute for Biotechnology (VIB), Katholieke
Universiteit Leuven, Herestraat 49, Leuven B-3000, Belgium
FEATURES
source
1.3566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p13-p12"
/issue_type="bone marrow"
1.3566
/gene="DUSP16"
/note="synonym: MKP7"
633..2630
/gene="DUSP16"
/EC_number="3.1.3.48"
/EC_number="3.1.3.16"
/codon_start=1
/product="dual specificity phosphatase 16"
/protein_id="AAW75120.1"
/translation="MAHEMIGTQIVTERLVALLSEGTREKVLIDSRPVEVNTSHILE
AININCSKLMKRLQODKVLITELIOHSAKHVDIDGOKVVYDSSODVASISDC
FLTVLLGKLEKSPNSVHLAGFAEFSPGELCRGKSTIVPTCSOPCLPVANIGPT
RLPNIYLGGCDVUNKELMOONGIYVLNANSTCPKPDFIPESFLVAVPNDSECEK
KRPITSPNENFLGQLDVEKIKNGTGAAGPSKIKLHLKPKNEPVPVAVEGQKSE
TPLSPPCADSATSEAGQRPVHPASVPSPVQPSLLEDSPVQALSGIHSADRIED
SKTKRSESLDISEVSAWMAASLHGSSSEDALEYKPSITLDGTNKLCOFSPVOE
LSEOTPEPSPDKSEASIPKLTQIARPSOSKRLHSVATSSGTAORSLSPLRSGS
VEDNHTSPRLGELTSCQHLTKSAGLGKGMHSDILAQOTSPSTSSWTATSSHP
YSASAIYGGASATYSCSGQLPTCGDYIVARRKQFSDRADSRSRWSHEEPFEKQFK
RSCQMEFGESIMSENRSREELGVSGSFSFGSMELIEVS"

Alignment Scores:
Pred. No.: 0 Length: 3566
Score: 472.00 Matches: 633
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
Gaps: 0

US-10-029-345A-109 (1-665) x AF506796 (1-3566)

Qy 1 MetAlahIsglMetIlglYthrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 633 ATGGCCCAATGATGATATTGGAATCAAAATTGTACTGAGAGGTGTGGCTGTGGA 692
Qy 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db 693 AGTGAACGGAAGAAAGTGTCTGCTAATTATATAGCCGCGCAATTGGAAATACATCATCC 752
Qy 41 HisIleleGluValAlaIleAsnIleAsnCySerSerIysMetIysArgArgLeuGln 60
Db 753 CACATTTTGAAGACCATTAATATCAATGCTCCAAAGCTTATGAAAGCGAAGTTGCAACAG 812
Qy 61 AspIysValLeuIleThrGluLeuIleGlnHisSerAlaValSHIlyValAspIleAsp 80
Db 813 GACAAAGTAAATTATACAGACTCATCCAGATTCAGCAAAACATAGGTGACATTGAT 872
Qy 81 CysSerGlnIysValIleValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 873 TCGAGTCAAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTCA 932
Qy 101 AspCysPheLeuThrValLeuLeuGluIlyLeuGluIlySerPheAsnSerValHisLeu 120
Db 933 GACTGTTTCTCAGTGTACTTCTGGGTAACTGAGAGAAAGACTTCAACTGTTCACCTG 992
Qy 121 LeuIleAglyGlyPheAlaGluPheSerArgCysPheProGlyLeuCySGluGlyIlySer 140
Db 993 CTGCAAGTGGGATTTGCTGAGATCTCTGTTGTTCCCTGCTGTGAAAGAAATTC 1052
Qy 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 1053 ACTCAGTCCCTACTGCTGATTTCTCAGCTTGTCTTACTGTCGCAATTGGGCCCAACC 1112
Qy 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIlyGluLeuIle 180
Db 1113 CGAATTTCTCCCAATCTTATCTTGTGCTGCAGAGATGTCCTCAAGAGACTGAT- 1171
Qy 181 -GlnGlnAspGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProIlySerProAsp 200
Db 1172 GCAGCAGATGGATGGATGTATGTATTAATGCCAGCA-TACTGTCCAAAGCTTACT 1230
Qy 200 heIleProGluSerHisPheLeuArgValProValaAsnAspSerPheCysGluIlySerIle 220
Db 1231 TTATCCCGGAGTCTCATTTCTGCTGTGCTGTGATGACAGTTTGTGAGAAATTT 1290
Qy 220 euPProTTPLeuAspIlySerValAspPheIleGluIlyValAlaIlyAsnIlyCysVal 240
Db 1291 TGCGGTGGTGGCAAAATCAGTAGATTTCATTGGAAGCAAAAGCTTCAATGAGATGTG 1350
Qy 240 alleuValHisCysLeuValAGIlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 260
Db 1351 TTCTAGTGCAGCTGTTAGCTGGGATCTCCGCTCCGCCACATCGCTATGCTTCAATCA 1410
Qy 260 eIlyValArgMetAspMetSerLeuAspGluAlaTyrArgPheValIlyGluIlyArgProT 280
Db 1411 TGAAGAGATGAGCATGTCTTTAGATGAACCTTACAGATTGTGAAAGAAAAAGACCTA 1470
Qy 280 hrIleSerProAsnPheAsnPheLeuGlyIleLeuLeuAspTyrGluIlyIlyIleIlyVal 300
Db 1471 CTATATCTCAAACTTCATATTTCTGGGCCAATCTCTGGACTATGAGAAATAATTAAAGA 1530
Qy 300 snGlnThrGlyAlaSerGlyProIlySerIlySerIlyLeuIlyLeuHisIleGluIlyPro 320
Db 1531 ACCAGACTGAGATCAGGCGCCAAAGACCAATCAAGCTGCTGCACTGAGAAAGCCAA 1590
Qy 320 snGluProValProAlaValSerGluGlyIlyGluIlySerSerGluThrProLeuSerPro 340
Db 1591 ATGAACCTGTCCTGCTGTCTCAGAGGTTGACAGAAAGCGAGCGCCCTCAGTCCAC 1650
Qy 340 rocCysAlaAspSerAlaThrSerGluAlaAlaIlyGlnArgProValHisProAlaSer 360
Db 1651 CTTGTGCCAGCTCTGCTACTCAGAGGACAGCAAGCAAAAGGCCCTGTGATCCGCGACGCG 1710

Qy	360	alProSeValProSeValGlnProSeLeuLeuGluAspSerProLeuValGlnAlaL	380
Db	1711	TGCCAGGCTGCCAGCTGCAGCCCTGCTGTTAGAGACAGCCGCTGGTACAGGGCC	1770
Qy	380	eusSerGlyLeuH1sLeuSerAlaApaGlyLeuGluAspSerAsnLySerLeuYsArSerP	400
Db	1771	TCAGTGGGCTGCACCTGTCCCGACACAGCTGGAGAGACAGCATATAGCTCAAGCTTCT	1830
Qy	400	heserLeuAspLleYsSerValSerTYrSerAlaSerMetAlaAlaSerLeuH1sGlyP	420
Db	1831	TCTCTCTGATATCAATCACTTTCATATTCAGCCAGCATGGCAGCATCTTACATGAGCT	1890
Qy	420	heserSerSerGluAspAlaLeuGluYrTYrLYsProSerThrThrLeuAspGlyTYrA	440
Db	1891	TCTCTCTATCAAGAAATGCTTTGGAAATCTAACAACTTTCATCTCTGATGGAGCA	1950
Qy	440	snLyLeuCYsGlnPheserProValGlnGluLeuSerGluGlnThrProGluThrSerP	460
Db	1951	ACAAAGCTATGCCAAGTCTTCCCTGTTCAGAACTTTCGAGACAGACTCCGAAACCAAGTC	2010
Qy	460	roAspLYsGluGluAlaSerLleProLYsLYsLeuGlnThrAlaArpProSerAspSerG	480
Db	2011	CTGATTAAGAGAGAGACGACGATCCCAAGAAAGCTGACAGCCGACGCTTCAAGCAAGCC	2070
Qy	480	InserLYsArGLeuH1sSerValArGThrSerSerSerGlyTYrAlaGlnArGserLeuL	500
Db	2071	AGACCAACCGATTGGATTGGTTCAGAACACACAGCAGAGTGGGACCGCCAGAGGTCCTTT	2130
Qy	500	eusSerProLeuH1sArGserGlySerValGluAspArTYrTh1sThSerPheLeuPheG	520
Db	2131	TATCTCCATGCTGATGAAAGTGGAGCGTGGAGAGACAAATTACACACCAAGCTTCTTTTCG	2190
Qy	520	lyLeuSerThrSerGlnGlnH1sLeuThrLYsSerAlaGlyLeuGlyLeuLYsGlyTYrP	540
Db	2191	GCTTTTCCACCAAGCAGCAGCAGCACTCAAGAAAGTCTGCGCTGGGCTTAAAGGCTTGGC	2250
Qy	540	IsSerAspLleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTYrTYrPhea	560
Db	2251	ACTGGGATATCTTGGCCCCCAAGCACTTACCCCTTCCCTACCGCAGCTGGTGAATTTTG	2310
Qy	560	laThrGluSerSerTh1sPheTYrSerAlaSerAlaLleTYrGlyGlySerAlaSerTYrs	580
Db	2311	CCACGAACTCTCACACTTCTACTCTGCTCAGCCATCTACGAGAGCGAGTCCAGTTACT	2370
Qy	580	erAlaTYrSerCYsSerGlnLeuProThrCYsGlyAspGlnValTYrSerValArGArG	600
Db	2371	CTGCTCAAGCTGCAGCAGCACTGCCCATTTCCGAGACCAAGTCTATTCTGTGGCAGGC	2430
Qy	600	rgGlnLYsProSerAspAlaAlaAspSerArGArGserTYrPheGlnGluSerProPheG	620
Db	2431	GGCAGAAAGCCAAAGTACAGAGCTGACTGGCGGAGCTGGCATGTAAAGAGCCCTTTTG	2490
Qy	620	lyLYsGlnPheLYsArGArGserCYsGlnMetGluPheGlyGluSerLleMetSerGluA	640
Db	2491	AAAAGCAATTTAAACGAGAAAGCTGCCCCAAATGGAAATTTGGAGAGAGCATCTATGTCAGAGA	2550
Qy	640	snArGSerArGlyGluLeuGlyLYsValGlySerGlnSerSerPheSerGlySerMetG	660
Db	2551	ACAGGTACCGGGAAGAGCTGGGGAAAGTGGGACATGACTTAAGCTTTTGGGACAGCATGG	2610
Qy	660	lyIleIleGluValSer 665	
Db	2611	AAATCATTTGAGGTCTCC 2627	
RESULT 15			
AX374994	AX374994	3766 bp	DNA linear PAT 01-MAR-20020
LOCUS	Sequence 17 from Patent W00210363.		
DEFINITION	AX374994		
ACCESSION	AX374994.1 GI:19169826		
VERSION			
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		

REFERENCE	Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Eumelostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 Tang,Y.T., Elliott,V.S., Ramkumar,J., Yao,M.G., Burford,N., Wang,Y.E., Stewart,E.A., Gandhi,A.R., Paterson,C., Lee,E.A., Hatfield,A.J., Lu,D.A., Triboley,C.M., Griffin,J.A., Baughn,M.R., Yue,H., Warren,B.A., Nguyen,D.B. and Walla,N.K.
TITLE	Protein phosphatase
JOURNAL	Patent: WO 0210363-A 17 07-FEB-2002;
FEATURES	Incyte Genomics, Inc. (US)
SOURCE	Location/Qualifiers 1..3766 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="Incyte ID No: 7480570CB1"
ORIGIN	
Alignment Scores:	
Pred. No.:	0 Length: 3766
Score:	472.00 Matches: 663
Percent Similarity:	99.40% Conservative: 0
Best Local Similarity:	99.40% Mismatches: 2
Query Match:	70.98% Indels: 4
Db:	6 Gaps: 0
US-10-029-345A-109 (1-665) x AX374994 (1-3766)	
QY	1 MetAlaHISgluMeTIIeGlyThrgInIleValIThrgIuArgLeuValAlaleuLeuGlu 20
Db	538 ATGGCCCATGAGATATATTGAACTCAAAATGTACTGAGAGGTGGTGGCTCTGCTGGAA 597
QY	21 SerGIYThrgIuLyValleuLeuIleAspSerARgProPheValGIuTYrAsnThrSer 40
Db	598 AGTGAAGCGAAAAAGTGTCTGTATTGATACCGGCCATTGTGTGAATACATACATCC 657
QY	41 HisIleuGluAlaAlIeAsnIleAsnCySerLySleMeTlyARgIArgLeuGIn 60
Db	658 CACATTTTGGAAAGCATTAATATCACTCAAGCTTCAGCTTATAGACGAAGGTTGCAACAG 717
QY	61 AspLySValleuIleThrgIuLeuIleGInIAsSerAlaIyHisIyValAspIleAsp 80
Db	718 GACAAAGGTTAATTACAGAGCTATCCAGCACTTACGCAACATRAGGTTGACATTGAT 777
QY	81 CySerGIuLyValIValIValTYrAspGInSerSerGIuAspValAlaserLeuSerSer 100
Db	778 TGCATCGAAGAGTTGTGATTTACATCAAGCTCCCAAGATGTTCCTCTCTCTTCA 837
QY	101 AspCyPheLeuThrValleuLeuGlyLySleuGIuLySerPheAsnSerValHisIleu 120
Db	838 GACTGTTTTTCACACTGTACTTCTGGGTAACTGGAGAGAGCTTCAACTCTGTTCACCTG 897
QY	121 LeuAlaGIyGlyPheAlaGluPheSerARgCyPheProGlyLeuCySgluGIyLySer 140
Db	898 CTTGAGAGGTGGTGTCTGAGTTCCTCTGTGTTCCTCGGCGCTCTGTGAAGAAAATCC 957
QY	141 ThrIleuValProThrCyAlIeSerGIuProCySleuProValAlaAsnIleGIyProThr 160
Db	958 ACTGTAGTCCCTACTCTGCAATTTCTAGCTTGCTTAACCTGTGTGCCAACATGGGCCAAC 101
QY	161 ArgIleuLeuProAsnIleuTYrLeuGlyCySgluARgAspValIleuAsnIySgluIle 180
Db	1018 CGAATTTTCCCAATCTTATCTTGGCTGCCAGCGAGAGTGTCTCAACAAGAGCTGAT- 107
QY	181 -GIuGIuAsnGIyIleGIyTYrValIleuAsnAlaserTYr-ThrCySPolyPProAsp 200
Db	1077 GCAGAGAGATGGAGATTGGTTATGTGTTAAATGCCAGCA-TRCCTGTCCMAAGCCGACT 113
QY	200 heIleProGluSerIAspHeuARgValProValAsnAspSerPheCySgluLyIle 220
Db	1136 TTATCCCGAAGTCTCATTTCTCTGCTGGCTGTGAATGACAGCTTTTGTGTGAANAATT 119
QY	220 euPTrIleuAspLySerValAspPheIleGluLyAlaLyAlaserAangIyCySv 240

```

Db      1196 TGCCGTGGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCATGAGATG 1255
Qy      240  alleuValHisyleuValIGlyIleSerArgSerAlaThrIleAlaIleAlaIleTyrIle 260
Db      1256 TTTAGAGGCACTGTTTAGTGGAGATCTCCCGCTCCGACCATCGCTATCGCTTACATCA 1315
Qy      260  etIysArgMetAspMetSerIleuAspGluAlaTyrArgPheValLysGluLysArgPro 280
Db      1316 TGAAGAGATGAGCATCTCTTAGATGACATTACGATTGTGGAAAGAAAAGAAAGACCT 1375
Qy      280  hrIleSerProAsnPheAsnPheIleuGlyIleuLeuLeuAspTyrGlyLysIleLys 300
Db      1376 CTATATCTCCAAACTTCAATTTTCTGGGCCAACTCCGACCTAGCACTGAGAGAAATTA 1435
Qy      300  snGIInthrGlyAlaSerGlyProLysSerIleuLysLeuLeuHisIleuGluLysPro 320
Db      1436 ACCAGACTGGAGATCAGGGGCCAAAGCAAACTCAAGCTGACCTGAGAAAGCCAA 1495
Qy      320  snGIuProValProAlaValSerGluGlyGlyIleLysSerGluThrProLeuSerPro 340
Db      1496 ATGAACTGTCTCTGCTGTCTCAGAGGTTGACAGAAAGCGAGACGCCCTTCAGTCCAC 1555
Qy      340  roCyAlaAspSerAlaThrSerGluAlaAlaGlyIleArgProValHisProAlaSerV 360
Db      1556 CCTGTGCCGACTCTGCTACTCAGAGGACAGAGCAAAAGGCCGCTGACATCCGCCAGCG 1615
Qy      360  alProSerValProSerValGlnProSerIleuLeuLysAspSerProLeuValGlnAla 380
Db      1616 TGCCCAAGCTGCCAGGAGTGACCCGCTGTGTAAAGACAGCCCGCTGGATCAGGCC 1675
Qy      380  euSerGlyLeuHisIleuSerAlaAspArgIleuLysAspSerAlaLysArgSerP 400
Db      1676 TCAGTGGGCTGCACCTGTCCGAGACAGGCTGGAACAGCAATAGCTCAAGCGTTCT 1735
Qy      400  heSerIleuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerIleuHis 420
Db      1736 TCTCTCTGATATCAATCAATCAGTTTCATATTACGCCAGCATGCGACATCTTACAT 1795
Qy      420  heSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrIleuAspGlyTh 440
Db      1796 TCTCCCTCATCAGAAAGTCTCTTGGATCTACCAAACTTCCACTCTCGATGGAGACA 1855
Qy      440  snLysLeuGlyGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSer 460
Db      1856 ACAAGCTATGCCAGTCTCCCGCTTCAAGAACTATCGAGAGACTCCCGAAACAGATC 1915
Qy      460  roAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSer 480
Db      1916 CTGATTAAGAGAGAGACCATCCCAAGAACTGACAGACCCCGCAGGCTTCAGACAGC 1975
Qy      480  InSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
Db      1976 AAGAGCAAGGATTCGATTCGTTCAGAACAGACAGAGTGAGCCCGCCAGAGATCCCTT 2035
Qy      500  euSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPhe 520
Db      2036 TATCTCACTGCATCGAAGTGGAGGTGAGAGACATTAACACACAGCTTCCTTTG 2095
Qy      520  LysSerThrSerGlnGlnHisIleuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrP 540
Db      2096 GCGTTTCCACGACGACGACGACCTCAGCAAGATCGCTGCGCTTAAAGGCGCTGGC 2155
Qy      540  IAspAspIleuAlaProGlnThrSerThrProSerLeuThrSerSerPyrPyrPhe 560
Db      2156 ACTGGATATCTTGGCCCCCGACACTTACCCCTTCCTGACACGACGCTGGATATTTG 2215
Qy      560  IAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyr 580
Db      2216 CCACAGAGCTCCACACTTCTACTCTGCTCAGCATCTACGAGAGGAGGTGCAAGTTACT 2275
Qy      580  eAlaIleTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600

```

```

Db      2276 CTGCCTACAGCTGCAGCCAGCTGCCACTTGGGAGACCAAGTCTTATTCTGCGCAGGC 2335
Qy      600  rGGLuLysProSerAspArgAlaAspSerArgSerTyrPheGluGluSerProPhe 620
Db      2336 GGCAGAGCGCAAGTACAGAGCTGACTCGCGCGAGCTGCGCATGAGAGAGCCCTTTG 2395
Qy      620  LuLysGluPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGlu 640
Db      2396 AAAGCAGATTTAACGCAAGAGCTGCCAAATGAAATTTGGAGAGAGCATATGTCAGAGA 2455
Qy      640  snArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetG 660
Db      2456 ACAGTTCAGGGGAAGAGCTGGGGAAAGTGGGACATGATCTAGCTTTTGGGACAGCAT 2515
Qy      660  LuIleIleGluValSer 665
Db      2516 AAATCATTTAGAGTCTCC 2532

RESULT 16
BD171157
LOCUS      BD171157          4790 bp    DNA          linear    PAT 17-JAN-2003
DEFINITION Novel gene and protein encoded thereby.
ACCESSION      BD171157
VERSION      BD171157.1  GI:27876969
KEYWORDS      WO 02052005-A/13.
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 4790)
AUTHORS      Ohara, O., Nagase, T. and Nakajima, D.
TITLE      Novel gene and protein encoded thereby
PARENT       WO 02052005-A 13 04-JUL-2002;
JOURNAL      KATUKA DNA RESEARCH INSTITUTE, OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE
NAKAJIMA

COMMENT
OS Homo sapiens (human)
PN WO 02052005-A/13
PD 04-JUL-2002
PF 20-DEC-2001 WO 2001JP011217
PR 22-DEC-2000 JP 00P 389742
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/12, C07K14/47
CC Novel gene and protein encoded thereby
FH Key location/Qualifiers
FT CDS location/Qualifiers
(184).. (2178).

FEATURES
source
1..4790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 4790
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
Gaps: 0
DB: 6

US-10-029-345A-109 (1-665) x BD171157 (1-4790)
Qy      1  MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
Db      184  ATGGCCCATAGATGATGGAACTCAAAATGTTACTGAGAGGTTGGCTCGCTGGAGAA 243
Qy      21  SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrPheThrSer 40
Db      244  AGTGAAGCAAGAAAAGTGGCTTAATTGATAGCCGCGCATTTGTGAAATACATATCATCC 303
Qy      41  HisIleuGluAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGlnGln 60
Db      304  CACATTTGGAAAGCCATTAAATATCACTGCTCCAAAGCTTATGAAGCGAAGTTGCAACAG 363

```

QY 61 AapLyVaIlleu1leThrg1leu1leg1nH1seSerAla1ySh1s1yVa1Aap1leap 80
 DB 364 GACAAAGTGTATTATTCAGAGCTCATCCAGATTCCAGAAACATTAAGTTGACATTGAT 423
 QY 81 CySeSerGln1yVa1Va1TyraApGlnSeSerGlnApyVa1Ase1leuSeSer 100
 DB 424 TGCAGTCAGAGGTGTAGTTTACGATCAAGAGTCCCAAGATGTTCCCTCTCTCTTCA 483
 QY 101 AapCySPhelEuThr1Val1leu1eug1y1y1eug1u1ySeSerPheAseSerVal1H1s1eu 120
 DB 484 GACTGTTTCTCACTGACTTCTGGGTAAACTCGAAGAGGCTTCAACTCTGTTCACTCG 543
 QY 121 leuA1aG1yG1yPhe1aG1uPheSeSerAryCySPhelEu1y1eug1y1eug1y1ySeSer 140
 DB 544 CTTCAGAGGTGGTTCGTAGTCTCTCGTGTTCCTGGCTCTGTGAAGAAATCC 603
 QY 141 Thr1euVa1ProThrCyS1leSeSerGlnProCyS1euProVa1A1aAn1leg1yProThr 160
 DB 604 ACTCTAGTCCCTACCTGCATTTCTCAGCCTTGCTTCTGTCACCAATGGGCCAACC 663
 QY 161 Arg1leuProAse1leu1y1eug1y1y1eug1AryApyVa1leuAse1y1eug1leu1le 180
 DB 664 CGAATCTTCCCAATCTTATCTTGGCTCGCAGCGAGATGCTCTCAACAAGAGCTGAT- 722
 QY 181 -GlnGlnaeng1y1leg1y1y1y1e1e1e1a1se1y1y1-ThrCyS1y1y1y1y1y1y1y1y1 200
 DB 723 GCAGCAAGATGGATGGTATGTATGTTAAATGCCACAA-TACTCTTCCAAAGCTGACT 781
 QY 200 he1leProGlnSeSer1H1sPhelEuAryVa1ProVa1AseApySePheCyS1y1y1le1 220
 DB 782 TTTATCCCGAGTCTATTTCTCGCTGCTGCTGTGAATGACAGCTTTTGTGAAGAAATTT 841
 QY 220 euProTr1euApy1ySeSerValApyPhe1leg1u1ySa1A1ySa1Ase1y1y1y1y1y1y1 240
 DB 842 TGCCTGGTGGCAATCAGTATGATTTCAATTGAGAAAGCAAAAGCCTCAATGATGATG 901
 QY 240 Al1euVa1H1s1y1e1e1a1y1le1Se1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 260
 DB 902 TTTCTAGTGAAGCTGTATGCTGGAGATCTCCGCTCCGACCAATCGCTATCGCTTACATA 961
 QY 260 e1y1 280
 DB 962 TGAAGAGATGACAGTCTTTTATGATGAAGCTTACAGATTTTGAAGAAAGAAAGCTTA 1021
 QY 280 he1leSe1y1 300
 DB 1022 CTATATCTCCAAACTTCAATTTTCTGGGCCAACTCTCGACATGAGAAAGATTAAGA 1081
 QY 300 eng1nThrg1y1Ase1y1 320
 DB 1082 ACCAGACTGAGAGATCAGGGCCAAAGAGCAAAAGCTGAGCTCGGAGAGAACCCAA 1141
 QY 320 eng1uProVa1ProA1aVa1Se1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 340
 DB 1142 ATGAACCTGTCCCTGCTGTCTCAGAGGTGACAGAAAGAGAGAGCCCTCAGTCCAC 1201
 QY 340 roCySa1AapSeSerA1aThrSe1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 360
 DB 1202 CCGTGCAGACTCTGCTACTCAGAGGACAGAGCAAAAGGCGGTGATCCCGCAGG 1261
 QY 360 a1ProSe1y1 380
 DB 1262 TGGCCAGCGTGGCCAGCGGACCGCTCGTGTAAAGACAGCCGCTGTGACAGGCC 1321
 QY 380 euSe1y1 400
 DB 1322 TCAAGTGGCTGACCTGTCCGACAGAGGCTGAAAGACAGCAATTAAGCTTCAAGCTTCT 1381
 QY 400 heSe1y1 420
 DB 1382 TCTCTGTGATATCAATCAGTTTCAATTAATTCAGCCAGCATGGAGCATCTTACATGGCT 1441

QY 420 heSeSeSerGlnApyA1a1e1u1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1 440
 DB 1442 TCTCTCATCAAGAGTGTGTTGAACTTACAAACCTTCCACTACTCTGATGGAGCA 1501
 QY 440 en1y1e1y1 460
 DB 1502 ACAAGCTATGCCAGTCTCCCTGTTCAAGAACTATCGAGAGAGACTCCCGAAACAGATC 1561
 QY 460 roApy1 480
 DB 1562 CTGATTAAG 1621
 QY 480 InSe1y1 500
 DB 1622 AAG 1681
 QY 500 euSe1y1 520
 DB 1682 TATCTTCACTGATGAG 1741
 QY 520 1y1e1y1 540
 DB 1742 GCTTTCCACAGCCAG 1801
 QY 540 1eSe1y1 560
 DB 1802 ACTGGATATCTTGGCCCCAGAGCTTACCTTCCCTTCCCTGACACAGAGAGAGAGAGAGAGAG 1861
 QY 560 1aThrg1y1 580
 DB 1862 CCAGAGAGCTTCACTCTTCACTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1921
 QY 580 e1y1 600
 DB 1922 CTGCTTACAGCTGAG 1981
 QY 600 1y1 620
 DB 1982 GGCAG 2041
 QY 620 1u1y1 640
 DB 2042 AAAG 2101
 QY 640 enArySe1y1 660
 DB 2102 ACAGGTCACGGAG 2161
 QY 660 1u1le1leg1y1 680
 DB 2162 AAATCATTAAGGCTTCC 2178
 RESULT 17
 BD183422
 LOCUS BD183422 4790 bp DNA linear PART 17-JUN-2003
 DEFINITION Novel genes and proteins encoded by the genes.
 ACCESSION BD183422
 VERSION BD183422.1 GI:31875622
 KEYWORDS JP 2002345492-A/135.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4790)
 AUTHORS Ohara, O., Nagase, T. and Nakajima, D.
 TITLE Novel genes and proteins encoded by the genes
 JOURNAL Patent: JP 2002345492-A 135 03-DEC-2002;
 KAZUSA DNA RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002345492-A/135
 PD 03-DEC-2002
 PF 26-FEB-2002 JP 2002049009
 PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA

PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00,
PC A61P25/14,
PC A61P25/18, A61P35/00, C12N15/00, A61K37/02
CC Novel genes and proteins encoded by the genes FH Key
Location/Qualifiers
FT CDS (184).. (2178).

FEATURES
source 1. .4790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 4790
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: Gaps: 0

US-10-029-345a-109 (1-665) x BD183422 (1-4790)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 184 ATGGCCCATGAGATGATGGAACTCAAAATGTTACTGAGAGGTGGTGGCTGCTGAGAA 243
QY 21 SerGlyThrGluLeuValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db 244 AGTGAACCGGAAAAGTGTGCTAAATGTAAGCCGCCCAATTTGTGAATACATATCATCC 303
QY 41 HisIleLeuGluAlaIleAsnIleAsnIleAsnIleSerLeuLeuMetLysArgLysLeuGln 60
Db 304 CACATTTTGGAAACCATTAATCAACTGCTCCAACTTATGAAAGCAAGTTGCAACG 363
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 364 GACAAAGTGTATTCAGAGCTCATCCAGCATTCAGCAAAACATAGGTTGACATTGAT 423
QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 424 TGCAGTCAGAAAGGTGTAGTTTACATCAAGATCCCAAGATGTGCTCTCTCTTCA 483
QY 101 AspCysPheLeuThrValLeuLeuGluLysLysLeuGluLysSerPheAsnSerValHisLeu 120
Db 484 GACTGTTTCTCACTGACTCTTCGGGTAACTGAGAAAGCTTCAACTCTGTTCACTG 543
QY 121 LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuCysGlnGlyLysSer 140
Db 544 CTTCAGGAGGTGGTGTGCTGAGTCTCTCGTTGTTCCCTGGCCTGTGAAAGAAATCC 603
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 604 ACTCTAGTCCCTACCTGCAATTTCTCAGCTTCTTAACCTGTTGCCAACTTGGGCAACC 663
QY 161 ArgIleLeuProAsnLeuTyrLeuGluCysGlnArgAspValLeuAsnLysGluLeuIle 180
Db 664 CGAATTCCTCCAACTTATCTTGGCTGCCAGCAGAGTCTCTCAACAGAGAGCTGAT- 722
QY 181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProLysProAsp 200
Db 723 GGAGCAGATGGATGGTATGTATGTAAATGCCAGAA-TACCTGTCCAAAGCTGACT 781
QY 200 HisIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
Db 782 TTATATCCGAGTCAATTTCTGCGGTGTGCTGTAAATGACGCTTTGTGAGAAATTT 841
QY 220 euProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysVal 240
Db 842 TGCCGTGGTGGCAAACTAGTATTTTCAATTTGAGAAAGCAAAAGCTCCCAATGATGTG 901
QY 240 allLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleLeu 260

Db 902 TTCTAGTGCACGTGTTAGTGGGATCTCCCGCTCCGCCACCATCGCTTACATCA 961
QY 260 eLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProT 280
Db 962 TGAAGAGATGACATGCTTATAGATGAAGCTTACAGATTTGTGAAGAAAAAGAACCTTA 1021
QY 280 HisIleSerProAsnPheAsnPheLeuGluGlnLeuLeuAspTyrCylLysLysIleLys 300
Db 1022 CTATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGATATGAGAAAGATTAAGA 1081
QY 300 SerGlnThrGlyAlaSerGlyProLysSerLysLysLeuLeuHisPheGluLysProAla 320
Db 1082 ACCAGACTGGAGCATCAGGCCCAAGACAAACTCAACTGCTGACCTGAGAAAGCCAA 1141
QY 320 SerGlnProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProp 340
Db 1142 ATGAACCTGCTCCCTGCTCTCAGAGGGTGAAGAAAGCGAGAGCGCCCTCAGTCCAC 1201
QY 340 roCysAlaAspSerLysPheThrSerGluAlaIleGlyGlnLysArgProValHisProAlaSer 360
Db 1202 CTGTGCGACTGTGTAACCTCAGAGGCGAGCAAGAAAGCCCGTGCATCCCGCAGCG 1261
QY 360 alProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaL 380
Db 1262 TGCCAGAGGTGCCAGCGTGCAGCCGTGCTTTAGAGAGACGCCCTGTTACAGGGCC 1321
QY 380 euSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSer 400
Db 1322 TCAGTGGGCTCACCCTGCCAGACAGCGCTGGAAGACAGCAATAGCTCAAGCGTTCT 1381
QY 400 hiserLeuAspLysLeuSerValSerTyrSerAlaSerMetAlaSerLeuHisGlyP 420
Db 1382 TCTCTCTGATATCAATCACTTCAATTTAGGACAGATGGACATCTTACATGGCT 1441
QY 420 hiserSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThra 440
Db 1442 TCTCTCATCAGAAAGTCTTGGATATCAAAACCTTCCACTGCTGATGGAGCA 1501
QY 440 snLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerP 460
Db 1502 ACAAGCTATGCAAGTCTCCCTGTTCAAGAACTATGAGAGCAACTCCGCAACCAAGTC 1561
QY 460 roAspLysGluGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSer 480
Db 1562 CTGATTAAGAGAGACCAAGCATCCCAAGAAAGCTTCAAGACCGGCTTCAACAGCC 1621
QY 480 InsertLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
Db 1622 AGAGCAAGCCATTCATTCGGTCAAGAACAGACAGCATGGGACCGCCAGAGTCCCTTT 1681
QY 500 euSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheG 520
Db 1682 TATCTCACTGCACTGAGAAAGTGGAGCGTGAAGCAATTCACACACAGCTTCTTTTCG 1741
QY 520 LysLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGluLysLeuGlyTyrP 540
Db 1742 GCTTTTCAACAGCAGCAGACAGCACTTCAAGAGTCTGCTGGCTTGAAGGCTGGCC 1801
QY 540 isSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheA 560
Db 1802 ACTCGATATCTTGGCCCCCAAGCTCTACCCCTTCCGTGACAGAGCTGGATTTTG 1861
QY 560 lathrGluSerSerHisPheTyrSerLysAlaSerAlaIleTyrGlyLysSerAlaSerTys 580
Db 1862 CCACAGAGTCTCACTTCTACTCTGCTCACAGCACTTACGAGGAGGAGGAGGAGTAACT 1921
QY 580 exAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
Db 1922 CTGCTTCAAGCTGACGACCTGCTCCACTTGGAGAGCAAGTATATTTGTGGAGGCG 1981
QY 600 rGlnLysProSerAspArgAlaAspSerArgArgSerTyrHisGluGluSerProPheG 620
Db 1982 GGCAGAGCCAAAGTGAAGAGCTGACTCGCGGGGAGGCTGCGATGAAGAGGCCCTTTG 2041

QY	620	IuLyvGlnPhnHyARgArGserCyGgiMnEtGluPhnGlyGluSerLeuMetSerGlnA	640
Db	2042	AAAAGCACTTTAAACGACGAAGCTGCCAATATGAAATTTGGAGAGACATCTGTCAAGA	2101
QY	640	sNArGserArGluGluNuGluLyvValGlySerGlnSerSerPheSerGlySerMetG	660
Db	2102	ACAGGTACGGGAAGAGCTGGGGAAGTGGGCAATCACTCACTTTTCGGGACGATCG	2161
QY	660	IuLeileIleGluValSer	665
Db	2162	AAATCATTTAGAGTCTCC	2178
RESULT 18			
AB051487	LOCUS	AB051487	4790 bp mRNA linear PRI 07-FEB-2001
DEFINITION		Homo sapiens mRNA for KIAA1700 protein, partial cds.	
ACCESSION		AB051487	
VERSION		AB051487.1	GI:12697944
KEYWORDS			
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE		Nagase,T., Kikuno,R., Hattori,A., Kondo,Y., Okumura,K. and Ohara,O.	
REFERENCE		Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro	
JOURNAL		DNA Res. 7 (6), 347-355 (2000)	
MEDLINE		21082932	
PUBMED		11214970	
AUTHORS		2 (bases 1 to 4790)	
TITLE		Ohara,O., Nagase,T. and Kikuno,R.	
JOURNAL		Direct Submission	
DEFINITION		Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,	
ACCESSION		Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba	
VERSION		293-0812, Japan (E-mail:cdna1nf@kazusa.or.jp.	
KEYWORDS		URI:http://www.kazusa.or.jp/huge, Tel:01-438-52-3913,	
SOURCE		Fax:01-438-52-3914)	
FEATURES			
source		Location/Qualifiers	
		1..4790	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="fj15353"	
		/note="vector:pb1uescriptII SK plus"	
		1..4790	
		/gene="KIAA1700"	
		<109..2181	
		/gene="KIAA1700"	
		/note="Start codon is not identified."	
		/codon_start=1	
		/product="KIAA1700 protein"	
		/protein_id="BAB21791.1"	
		/db_xref="GI:12697945"	
		/translation="APEHTTTSSMOIKEGGKRGILVMAHMIQIOTERIVALL	
		ESGTEKULIRSRPEVNTSHITTEALININSKLMKRLQDDKYLITLTHSRAKHY	
		DIDCSQKVVYDSSQDVASISDCEFTLVLGKLEKSFNSVHLAGFAEIRFCPEL	
		TCGPSPDIPFICISQPCLEPVANIGPTRLPLNYLGCQDQVNLKELMOONGIGYVLA	
		ATIAIAYIMKMDMSLDEAYAFVKEKEPTISPNPNFGOLLDYDKIKINOTGASGPG	
		KLILHLEKPEPEPVAYSEGQKRETPISPCACASITGSAAGORPVHAPVSPVSPQ	
		PSLEDESLINQALSGILHSLAPRLDSNKLKRSFELDIKSYISMAASLHGRSSSD	
		ALHYKPESTTLIDGTRNKLQCSFVVELSQQPEIISPDKEASILPKLQTPARPSDSQR	
		LHSVPTSSSGTAAGSLSPLHRSGLSEVDNNTYSLFLGLSTGQQLHTLSAGLKGWMS	
		DILAPQSTPSLTSWYFATSSHSYFSAIYCGSASYSAVSCQLPTCGDOVYSVRR	
		ROKSDRADSRSSRWHEESPFEKQFKRSQVCEMFESINSENRSELKYGQSSFGSG	
		SMEIIEVS"	
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	4790

Score:	472.00	Matches:	663
Percent Similarity:	99.40%	Conservative:	2
Best Local Similarity:	99.40%	Mismatch:	0
Query Match:	70.98%	Indels:	4
DB:	9	Gaps:	0
US-10-029-345A-109 (1-665) × AB051487 (1-4790)			
QY	1	MetaLahisglumetiIegIYThrcniIleValThrguArlgleuValAlaleuLeuglu	20
Db	184	ATGGCCCATGAGATGATGTAAGAACCTCAAAATTGTTACTGAGAGTTGGTGGCTGCGTGA	243
QY	21	SerGIYThrguIyValleuLeuIleAspSerArlProPheValGIuTYrAenThser	40
Db	244	AGTGAAGGAAAAAGTCTCTAATTATGACCGGCATTGTGTGAATACAAATCAATCC	303
QY	41	HisIleIeugIuAlIeAsnIleAsnCySerIySleuMetIyArlArgLeuIngIn	60
Db	304	CACATTTTGGAAAGCCATTAAATTCAACTGCTCCAGACTTAAGAACCGAAGTTGCACAG	363
QY	61	AspIyValIeuleIeThrguIeuleIleGlnHisSerAlalyShIeIyValAspIleAsp	80
Db	364	GACAAAGCTTAATTACGAGCTCATCCAGACTTCAGGAAACAAATAGGTTGACATTGAT	423
QY	81	CysSerGIuIyValValIyTYrAspGlnSerSerGlnAspValAlIeSerIeUser	100
Db	424	TGCAGTCAGAAAGTTGTGATTACCATCAAAAGCTCCAAAGATGTTGCCCTCTCTCTTCA	483
QY	101	AspCySPheIeUThrValIeUleuGIyIyIeugIyIySerPheAsnSerValHisIleU	120
Db	484	GACTGTTTTCACACTGACTCTGGGTAACATGAGAAAGACTTCACTGTGTACCTG	543
QY	121	LeuAlaGIyGIyPheAlaGIuPheSerArgCySPheProGIyIeUCyGIuIyIySer	140
Db	544	CTTGAGAGTGGGTTGCTGAGTGTCTCTGTTGTTCCCTGGCTCTGTGAAGAAATCC	603
QY	141	ThrlleuValProThrCySIIeSerGlnProCySleuProValAlAlaSnIleGIyProThr	160
Db	604	ACTCTAGTCCCTACCTGCATTTCTCAGCCTTGCTTACCTGTGTCACAAATGGGCCAAC	663
QY	161	ArgIleIeUProAsnIeUThrIeUGIyCySgInArlAspValIleuAsnIyGIuIeUle	180
Db	664	CGAATCTTCCCAATCTTAATCTTGGCTGCCAGCAGATGTCCTCAACAGAGAGCTGAT	722
QY	181	GlnGlnAsnGIyIleGIyTYrValIleuAsnAlaSerTYr-ThrCySProlysProAsp	200
Db	723	GCAGCAGAAATGGGATGGTTATGTCTTAAATGCCAGCAA-TACCTGTCCAAAGCCTGACT	781
QY	200	heIleProGIuSerHisIePheIeUArGValProValAsnAspSerPheCyGIuIyIleU	220
Db	782	TTATCCCGAGATCATTTCTCGGGGTGGCGCTGATAGACAGCTTTTGTGAAATAATT	841
QY	220	eUProTlPLeuAspIySerValAspPheIleGIuIyAlaIyAlaSerAsnGIyCyGv	240
Db	842	TGCCCTGGTTGACAAATCAGTAGATTTCATTGAAAGAACAAACCTCCCATGTGATGG	901
QY	240	alleuValHisCyIeUeUAlaGIyIleSerArgSerAlaThrIleAlaIleAlaTYrIleM	260
Db	902	TTCTTAGTCACTGTTTACTGGGATCTCCGCTCCGCCAACATGCTTATCCGCTCATCA	961
QY	260	eUlySArgMetAspMetSerIeUAspGIuAlaTYrArgPheValIySgIuIySArgProT	280
Db	962	TGAAGAGATGACATGTCTTATAGTAGAAGCTTACAGATTGTGTGAAGAAAAAGACCTA	1022
QY	280	hrlIeSerProAsnPheAsnPheIeUGIyGlnIleUAspTYrGIuIySlyIleIyBA	300
Db	1022	CTAATATCCCAAACCTCAATTTTCTGGGCCCAACTCTGTGACTATGAGAAAGATTAAGA	1081
QY	300	enGIInThrGIAlaSerGIyProIySerIySleuIySleuIySleuIySleuGIyIyProA	320
Db	1082	ACCAAGACTGGAGCATCAGGCCCAAGAACCAACTCAAGCTGCTGCACTGTGAGAGGCCAA	1144
QY	320	enGIUProValProAlaValSerGIuGIyGIyGlnIyIySerGIuThrProIeUserProp	340


```

Db      1142 ATGAACCTGTCCTGCTGTCTCAGAGGCTGACAGAAAACGAGAGCCCTCAGTCCAC 1201
Qy      340  rocyssalaaasppseralathsergluaalaaaglylnaagp-rovallhisproalaser 360
Db      1202 CCTGTGCGACTGCTGTAACCTCAGAGGACAGACAGAAAGCCCGCTGCACTCCGCGACAGG 1261
Qy      360  alProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlu 380
Db      1262 TGCCACAGCGTCCAGGCTGACACCGCTGCTTGAAGAGACACCCGCTGTACAGGCGC 1321
Qy      380  euSerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAnlySLeuYsArgSerp 400
Db      1322 TCAGTGGGCTGCACCTGTCCGACAGAGCTGGAGACAGCATTAAGCTCAAGCGTTCTCT 1381
Qy      400  heSerLeuAspLileysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyP 420
Db      1382 TCTCTCTGATATCAATCAATCAAGTTTCATATTCACCCAGCATGGACATCTTACATGGCT 1441
Qy      420  heSerSerSerGluAspAlaLeuGlnUlyTyrTyrPProSerThrThrLeuAspGlyThra 440
Db      1442 TCTCCATCAGAAAGATGCTTGGATATCTACAAACCTTCCACTCTGTGATGGACCA 1501
Qy      440  snlySLeuCySglnPheSerProValGlnGlnLeuSerGlnGlnThrProGluThrSerp 460
Db      1502 ACAAGCTATGCCAGTTCTCCCTGTTCAGGAACTATCGAGAGCAGACTCCCGAAACAGTC 1561
Qy      460  roAspLySglnGlnUalaserLileProLySlyLeuGlnThrAlaAspProSerAspSerG 480
Db      1562 CTGATTAAGAGAGAGCAGATCCCAAGAAAGCTGACAGACCGCAGAGCTTTCAGACAGGC 1621
Qy      480  InSerLySArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuL 500
Db      1622 AGAGCAAGCGATGTCATTTGGTCAGAAACAGACAGCGATGGACCGCCAGAGAGTCCCTTT 1681
Qy      500  euSerProLeuHisArgSerGlySerValGluAspAntYrHisThrSerPheLeuPheG 520
Db      1682 TATCTCCTACCTGATTCAGAGTGGAGGCTGAGAGACATTAACACACACCTTCTTTTCG 1741
Qy      520  lYLeuSerThrSerGlnGlnHisLeuThrLySeraLaglyLeuGlyLeuLySgLYTPH 540
Db      1742 GCTTTTCCACAGCCAGCCAGCACCTTCAGAAAGTCTGCTGGCCCTTAAGGCGTGGC 1801
Qy      540  iSerAspLileuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheA 560
Db      1802 ACTCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGAGCTGTGATTTTG 1861
Qy      560  laThrGluSerSerHisPheTyrSerAlaSerAlaLileTyrGlySerAlaSerTyrS 580
Db      1862 CCACAGAGTCTTCACACTTCTACTCTGCTCAGCCATTCACGAGGAGTGCACATTTACT 1921
Qy      580  eAlaTyrSerCySerGlnLeuProThrCySgLYAspGlnValTyrSerValArgArga 600
Db      1922 CTGGCTACAGCTGCAGCCAGCTGCCCTTGGCGAGAACCAAGTCTTAATCTGTGCGCAAGC 1981
Qy      600  xgGlnLySProSerAspArgAlaAspSerArgArgSerTyrHisGlnGluSerProPheG 620
Db      1982 GGCAGAGAGCCAGAGTACAGAGCTGACTGCGCGGCGAGCTGGCATGAAGAAGGCCCTTTG 2041
Qy      620  lnlYglnPheLySArgArgSerCySglnMetGlnPheGlyGlnSerLileMetSerGluA 640
Db      2042 AAAACAGATTAAACGACAGAAAGCTCCCAATGGAAATTTGGAGAGGACATCATGTCAAGA 2101
Qy      640  snArgSerArgGlnGluLeuGlyLySValGlySerGlnSerSerPheSerGlySerMetG 660
Db      2102 ACAAGCTACGGGAAAGAGCTGGGAAAGTGGGACATCATGTCTTTCGCGGACACATGG 2161
Qy      660  lnllellleGlnValSer 665
Db      2162 AAATCATGAGGTCCTC 2178

```

```

LOCUS      AX099933          2118 bp    DNA          linear      PAT 02-APR-2001
DEFINITION Sequence 15 from Patent WO0120004.
ACCESSION  AX099933
VERSION     AX099933.1   GI:13538943
KEYWORDS   SOURCE
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE   1
AUTHORS    Yue,H., Tang,Y.T., Bandman,O., Hillman,J.L., Baughn,M.R.,
            Azimzai,Y. and Lu,D.A.
TITLE       Protein phosphatase and kinase proteins
JOURNAL     Patent: WO 0120004-A 15 22-MAR-2001;
            Incyte Genomics, Inc. (US)
FEATURES   source
            1..2118
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            /note="Incyte ID NO: 1234795CB1"

ORIGIN
Alignment Scores:
Pred. No.:      0          Length:      2118
Score:          439.00     Matches:      663
Percent Similarity: 98.96%   Conservative: 0
Best Local Similarity: 98.96%   Mismatches: 2
Query Match:    66.02%       Indels: 7
DB:              Gaps:      0

US-10-029-345A-109 (1-665) x AX099933 (1-2118)
Qy      1  MetAlaHisGluMetLileGlyThrGlnLileValThrGluArgLeuValAlaLeuLeuGln 20
Db      25  ATGGCCCATGAGATGATGGAACTCAAAATGTTTCTCTAGAGAGTTGGTGGCTCTGCTGAA 84
Qy      21  SerGlyThrGlnLySValLileLeuLileAspSerArgProPheValGlyTyrAspThrSer 40
Db      85  AGTGGACGAAAGAAAGTCTCTTAATGATAGCCGCAATTTGGAAATACAAATACATCC 144
Qy      41  HisLileuGlnAlaLileasnLileasnCySerLySLeuMetLySArgArgLeuGln 60
Db      145  CACATTTTGGAAAGCATTAATATCACTGCTCCAAAGCTTATGAAGGAGTTTGCAACAG 204
Qy      61  AspLySValLileuLileThrGluLeuLileGlnHisSerAlaLySHisLySValAspLileAsp 80
Db      205  GACAAAGTGTAAATTACAGAGCTCATCAGCATTCACCGAAACATTAAGTTGACATTGAT 264
Qy      81  CySerGlnLySValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db      265  TGCAGTCAGAGAGTTGATGATTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 324
Qy      101  AspCySphLeuThrValLileuLeuGlyLySLeuGlnLySLeuPheAsnSerValHisLeu 120
Db      325  GACTGTTTTCTCACTGTAATCTTGGGTAAACTGGAGAAAGCTTCAACTGTTCACCTG 384
Qy      121  LeuAlaGlyGlyPheAlaGlnPheSerArgCySphProGlyLeuCySglnGlyLySLeu 140
Db      385  CTTCAGAGTGGGTTTGTCTGATGTTCTTCGTTGTTTCCCTGCTGCTGGAAGAAATCC 444
Qy      141  ThrLeuValProThrCySHisSerGlnProCySLeuProValAlaAsnLileGlyProThr 160
Db      445  ACTCTATCTCCCTACCTGATCTTCTCAGCTTGTCTTACTTGTCCAAATTTGGGCAACC 504
Qy      161  ArgLileuProAsnLeuTyrLeuGlyCySglnArgAspValLeuAsnLySglnLeuLile 180
Db      505  CGAATCTCTCCCAATCTTATCTTGGCTGCCACAGAGATGCTCTCAACAAAGAGCTGAT- 563
Qy      181  -GlnGlnAsnGlyTlileGlyTyrValLeuAsnAlaSerTyr-ThrCySProLySProAsp 200
Db      564  GCAGCAAGATGGAGATTGGTTATGTGTTAAATGCAAGCA-TACTCTGTCCAAAGCTGACT 622

```


QY 200 heileProgluSerHisPheLeuArgValProValAsnAspSerPheCysGluValIle 220
 DB 623 TTTATCCCCGAGTTCATTTTCTGCTGCTGCTGATGACACTTTTGTGAGAAAATTT 682
 QY 220 euProTrpLeuAspLysSerValAspPheIleGluValAlaValAspAsnGlyCysV 240
 DB 683 TGCCGATGTTGACCAATCATGATGATTTCTATGAAAGCAAAAGCTCCAAATGATG 742
 QY 240 aileuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleATyrIle 260
 DB 743 TTTCTATGACACTGTTTATGCTGGGATCTCCGCTCCGACCATCGCTATCGCTTACATCA 802
 QY 260 euLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluValArgPro 280
 DB 803 TGAAGAGATGACATGTTCTTTAGATGAACTTACAGATTTGTGAAGAAAAGACCTA 862
 QY 280 hrtIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluValLysAla 300
 DB 863 CTATATCTCCAACTTCAATTTTCTGGGCCAATCTCGACCTATGAGAAAGATTAA 922
 QY 300 enGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisGluLysPro 320
 DB 923 ACCAGACTGGAGCATCAGGGCCAAAGCAAGCTGCTGACCTGGAGAAACCA 982
 QY 320 enGluProValProAlaValSerGluGlyGlnLysSerGluThrProLysSerPro 340
 DB 983 ATAAACCTGTCCTGCTGCTGCTGACAGAGGTGACAGAAAGCGAGAGCCCTCAGTCCAC 1042
 QY 340 roCysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSer 360
 DB 1043 CCGTGCCGACCTGCTGCTGCTGCTGACAGAGGCAAGAAAGCCCGTGACCTCCGCAAG 1102
 QY 360 aProSerValProSerValGlnProSerLeuGluLysAspProLeuValGlnAla 380
 DB 1103 TGCCGACGCTGCCAGGTGACAGCCGCTGTTAGAGGACAGCCGCTGTTACAGGCG 1162
 QY 380 euSerGlyLeuHisLeuSerAlaAspArgLeuGluLysAspAsnLysLeuValArgSer 400
 DB 1163 TCGTGGGCTGACCTGCTGCTGCTGACAGAGGCTGAGAAAGCAATTAAGCTTAAGCTTCC 1222
 QY 400 heSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGly 420
 DB 1223 TCTCTCTGATATCAATCATGTTCTTCAATTCACCCAGCAGCAGCATCTTTCATGAGCT 1282
 QY 420 heSerSerSerGluAspAlaLeuGluLysTyrLysProSerThrThrLeuAspGlyThr 440
 DB 1283 TCTCTCATCATGAAAGATGCTTGGAAATACAAACCTTCCACTACTGATGGAGCA 1342
 QY 440 enLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSer 459
 DB 1343 ACAAGCTATGCGCAATTCCTCCCTGTTCAAGAACTATCGAGCAGACTCCGAAACCAAGTT 1402
 QY 460 -ProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAsp 479
 DB 1403 CCTGATTAAGAGAGAGAGCCAGCATCCCAAGAAAGCTGACAGCCAGAGCTTCAAGAC 1462
 QY 479 rgLysLeuValArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSer 499
 DB 1463 CCAGACCAAGCGATTCATTCGCTCAGAACACAGACAGAGTGCGCCAGAGAGTCCCT 1522
 QY 499 uLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeu 519
 DB 1523 TTTATCTCCACTCATCGAAGTGGAGCGTGGAGGCAATTAACACCGACTTCCCTTT 1582
 QY 519 eglLysSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysLeuGlyT 539
 DB 1583 CGGCTTTTCCACCAAGCAGACAGACACTCAGAAAGTCTGCTGGCTGGCCCTTAAGGCT 1642
 QY 539 rpiHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrTyr 559
 DB 1643 GGCACCTCGAATCTTGGCCCCCAGACCTTACCCCTTCCAGACAGAGCTGATAT 1702
 QY 559 heAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSer 579

DB 1703 TTGCCAGACAGATCTCTACACTTCTACTGCTGCTGACCATCTTACAGAGGCAAGGCCAGTT 1762
 QY 579 yrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArg 599
 DB 1763 ACTGTGCTTACAGCTGACAGCCAGCTGCTCCACTTGGCGAAGCAAGCTTATCTGTGCGCA 1822
 QY 599 rGArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSerPro 619
 DB 1823 GCGCGCAGAAAGCCAGTGCACAGAGCTGACTCCGCGGAGCTGGCARTGAAGAGGCCCT 1882
 QY 619 heGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSer 639
 DB 1883 TTGAAGAGACATTTTAACCCAGACAGCTGCCAAATGAAATTTGGAGAGACATCATCTCAG 1942
 QY 639 lAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerPheSerGlySer 659
 DB 1943 AGAACAGCTCAGCGGAAGAGCTGGGGAAGTGGGAGTCACTAGTCTTTCGCGCAGCA 2002
 QY 659 eGluIleIleGluValSer 665
 DB 2003 TGAATATCATGAGGCTTCC 2022
 RESULT 20
 AX921915 2200 bp DNA linear PART 18-DEC-2003
 LOCUS AX921915
 DEFINITION Sequence 255 from Patent WO02068649.
 ACCESSION AX921915
 VERSION AX921915.1 GI:40215409
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 JOURNAL Patent: WO 02068649-A 255 06-SRP-2002;
 CURATOR Curation Corporation (US)
 FEATURES
 source location/Qualifiers
 1..2200
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2200
 Score: 408.00 Matches: 435
 Percent Similarity: 99.328 Conservative: 0
 Best Local Similarity: 99.328 Mismatches: 0
 Query Match: 61.358 Indels: 3
 DB: 6 Gaps: 0
 US-10-029-345A-109 (1-665) x AX921915 (1-2200)
 QY 231 GIULYBALALYALASERXNGLYCYSVALLLEUVALHICYLEUVALAGLYLIESERARG 250
 DB 793 GAGAAAGCAAAAGCCCAATGATGTGTTCAAGTGCACCTGTTTACCTGGAGATCTCCCG 852
 QY 251 SerAlaThrIleAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAla 270
 DB 853 TCGGCACCATGCTATCGCTTACATCATGAAAGAGATGACATGCTCTTACATGAAGCT 912
 QY 271 Tyr--ArgPheValLysGluLysArgProThrIleSerProAsnPheAsnPheLeuGly 289
 DB 913 TACAGAGATTTGTGAAGAAAAGACCTACATATCTCAAACTTCAATTTTCTGGGCG 972
 QY 290 GlnLeuLeuAspTyrGluLysValLysValLysValLysValLysValLysValLysVal 309
 DB 973 CAATCTCGAATATGAGAAAGAAATTAAGAAACAGACTGAGCACTCAGGCGCAAGAGAC 1032
 QY 310 LysLeuLysLeuLeuHisLeuGluLysProAsnGluProAlaValSerGluGly 329

Db	1033	AAACTCAGCTGTGTGCACTCGAGAAAGCAAAATGAACTGTCTCCGTGTCTCAGAGGCT	1092
QY	330	GIYGLINLYSSErGIunThrProIeuSerProProCYsAlaASPSeRAlaThrSerGIunAla	349
Db	1093	GGACAGAAAAAGCAGACGCCCTCAGTCCAGCTCGAGCTTGCCGACTTGCTACCTCAGAGGCA	1152
QY	350	AlaGIYGLINaTrProValHisProAlaSerValProSerValProSerValGlnProSer	369
Db	1153	GCAGGACAAAGGCCCTGTGCATCCCGCACGCTGCCAGGCTGCCAGGCTGCAGCCGTGC	1212
QY	370	LeuLeuGLIaASPSeRProIeuValGlnAlaIeuSerGIYleuHisIeuSerAlaAspArg	389
Db	1213	CTGTTAGAGACAGAGCCCGCTGGTACAGGCGCTCAGTGGAGCTGACCTGTCCGACAGAGG	1272
QY	390	LeuGIuASPSeRAsnLYSleuLYsArgSerPheSerIeuAspLIleLYsSerValSerTYr	409
Db	1273	CTGGAAAGACAGCAATTAAGCTCAAGGGTCTCTCTCGAATATCAATTCAGTTTCATAT	1332
QY	410	SerAlaSerMetAlaAlaSerIeuHisGlyPheSerSerSerGIuASPAlaIeuGIunTYr	429
Db	1333	TCAGCGACAGATGGACAGCACTCCCTTAATGGCTTCTCTCATCAGAAAGCTTTGGAAATAC	1392
QY	430	TYrLYSProSeRThrThrIeuAspGIYThrAsnLYSleuCYsGlnPheSerProValGln	449
Db	1393	TACAAACCTTCCACTACTCTGTGATGGAGACCAACAGACTTATGCCATTTCTCCCTGTTACG	1452
QY	450	GIuIeuSerGIunGlnThrProGIunThrSerProAspLYSGIuAlaSerIleProLYs	469
Db	1453	GAACATATGGAGACAGACTCCGAAACCAAGTCTGTATAGAGAAACAGCATCTCCCAAG	1512
QY	470	LYSleuGlnThrAlaArgProSerAspSerGlnSerLYsArgIeuHisSerValArgThr	489
Db	1513	AACTGCAACACCGCCAGGGCTTCACACAGCCAGAACAGCATTCGATTCGGTCAGAAC	1572
QY	490	SerSerSerGIYThrAlaGlnArgSerIeuIeuSerProIeuHisArgSerGIYSerVal	509
Db	1573	AGCAGCAGTGGACACGCCGCCAGAGGCTCCCTTTATATCCACTGCACATCGAAAGGGAGGGTG	1632
QY	510	GIuAspAsnTYrHisThrSerPheIeuPheGIYleuSerThrSerGIunGlnHisIleuThr	529
Db	1633	GAGGACAATTATCACACACGAGTCTCTTTTGGCCCTTCCACACGACGACAGCACTCAG	1692
QY	530	LYSserAlaGIYleuGIYleuLYSGLYTrpHisSerAspLIleuAlaProGlnThrSer	549
Db	1693	AAGTGTGCTGGCTGGGCTTAAGGGCTGGCACTCGCATATCTTTGGCCCCCAGACCTCT	1752
QY	550	ThrProSerIeuThrSerSerTrpLYrPheAlaThrArgIeuSerHisPheTYrSerAla	569
Db	1753	ACCCTTCCCTGACCAAGACAGCTGGATTTTGTGCACAGAGTCTTCAACTTACTTCTGACC	1812
QY	570	SerAlaIleTYrGIYGIYSerAlaSerTYrSerAlaTYrSerCYsSerGlnIeuProThr	589
Db	1813	TCAGGCATCTACGGAGGCGAGTGCAGATTACTCTGCTACAGCTGCACAGCTGCCACT	1872
QY	590	CysGIYAspGlnValTYrSerValArgArgArgGlnLYSProSerAspArgAlaAspSer	609
Db	1873	TGCGGAGACCAAGTCTATTCTGTGTGGCGAGCGGACAGAAAGCCAAAGTGCACAGAGCTGC	1932
QY	610	ArgArgSerTrpHisGlnGlnIeuSerProPheGlnLYSGlnPheLYsArgArgSerCYsGln	629
Db	1933	CGGCGAGACTGCATTAAGAGAGGCCCTTTGAAAGACGTTTAAACCCAGAAAGCTGCCAA	1992
QY	630	MetGlnPheGIYGIuSerIleMetSerGIuAsnArgSerArgGIuGlnIeuGIYLYsVal	649
Db	1993	ATGGAAATTGGAGAGAGCATATGTCAAGAGAACAGGTCAAGGAAAGAGCTGGGAAAGTG	2052
QY	650	GIYserGlnSerSerPheSerSerGIYSerMetGlnIleIleGlnValSer	665
Db	2053	GGCAGTCAGTCTAGCTTTTCCGGGACAGCATGGAATCATTTAGAGGTCTCC	2100

DEFINITION	Homo sapiens 12 BAC RP11-253119 (Roswell Park Cancer Institute Human BAC library) complete sequence.
ACCESSION	AC007619
VERSION	AC007619.23
KEYWORDS	GI:25075851
SOURCE	HTC
ORGANISM	Homo sapiens (human)
AUTHORS	Human sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 172206) Munzly,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Aladrooks,S.L., Amaratunge,H.C., Aye,J.R., Ayelle,M., Banks,T., Barbatta,R.J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva-M., Brown,E., Brown,M., Bryant,N.P., Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chin,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathnonn,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Deederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Donthaite,K.J., Draper,H., Dugan-Rocha,S., Dublin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emeling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P., Gabisi,A., Gao,Y., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,U., Harris,K., Harris,K., Hart,M., Havlik,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Honesi,F., Howard,S., Huber,J., Hulik,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivest,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Koravaj,J., Kovac,C., Kraticovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,U., Li,Z., Lichtarge,O., Liu,C., Liu,W., Louisseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mahoney,E., McLeod,M.P., Meador,M., Mel,G., Mescher,T., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Moebacht,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nicholson,E., Nockenkwu,S., Ogun,M., Okwuonu,G., Oraguene,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primm,B., Pu,L.L., Quiles,M., Ran,Y., Rivers,M., Rojias,A., Rojibokem,I., Rolfe,M., Ruiz,S., Savery,I., Scherer,S., Scott,G., Shen,H., Shm,C., Shooshtarian,N., Sisson,I., Sodereyen,E., Sonake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabot,P., Tameits,A., Tameits,K., Tang,H., Tansey,U., Taylor,C., Taylor,T., Teitford,B., Thomas,N., Thomas,S., Umanai,K., Vazquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Noore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 172206)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (21-MAY-1999) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 172206)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (28-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	4 (bases 1 to 172206)
AUTHORS	Morley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (03-MAY-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	5 (bases 1 to 172206)
AUTHORS	Morley,K.C.

TITLE Direct Submission
JOURNAL Submitted (07-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 6 (bases 1 to 172206)
AUTHORS Worley, K.C.
TITLE Direct Submission
JOURNAL Submitted (28-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 7 (bases 1 to 172206)
AUTHORS Worley, K.C.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Nov 19, 2002 this sequence version replaced gi:7658297.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES

source Location/Qualifiers
1..172206
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-253119"
176..413
/standard_name="G41293"
complement(888..1190)
/rpt_family="AluSx"
repeat_region
complement(1675..1981)
/rpt_family="AluYb8"
repeat_region
complement(1985..2232)
/rpt_family="AluSx"
repeat_region
complement(2398..2437)
/rpt_family="AluSx"
3010..3016
/function="low quality"

repeat_region 3300..4444 /rpt_family="L1MA8"
repeat_region 4449..4749 /rpt_family="AluUo"
repeat_region 4880..5143 /rpt_family="L2"
repeat_region 5192..5266 /rpt_family="L2"
repeat_region complement(6255..6342)
/rpt_family="L1PA7"
repeat_region 7602..7623 /rpt_family="AT_rich"
repeat_region 8096..8126 /rpt_family="(TGS)n"
repeat_region complement(8135..8267)
/rpt_family="AluUo"
repeat_region 8296..8586 /rpt_family="AluSp"
repeat_region 8911..8938 /rpt_family="AT_rich"
complement(10581..10724)
/rpt_family="MIR"
repeat_region 10995..11290 /rpt_family="AluUo"
repeat_region 14543..14574 /rpt_family="AT_rich"
complement(15322..15625)
/rpt_family="AluSx"
repeat_region 17554..17716 /rpt_family="FRAM"
repeat_region 17785..18079 /rpt_family="AluSx"
repeat_region 18231..18527 /rpt_family="AluSp"
repeat_region 18706..18741 /rpt_family="AT_rich"
repeat_region 18901..19190 /rpt_family="AluUo"
repeat_region 19379..19754 /rpt_family="MIR"
complement(20289..20337)
/rpt_family="MIR"
repeat_region 21165..21658 /rpt_family="L1MA9"
complement(22241..22463)
/rpt_family="L2"
repeat_region 22638..23279 /rpt_family="L2"
complement(24110..24233)
/rpt_family="L2"
STS 24235..24361 /standard_name="G62270"
STS 24267..24404

Alignment Scores:

Pred. No.: 0 Length: 172206
Score: 394.00 Matches: 394
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.25% Indels: 0
DB: 9 Gaps: 0

US-10-029-345A-109 (1-665) x AC007619 (1-172206)

QY 272 ATGPhEValIyGULyLArGrProThriIleSeProAenPheAenPheLeuGlyGlnLeu 291
|||||
DB 55841 AGATTTCGAAAGAAAAGACCTATATCTCCAACTTCAATTTCTGGGCCAATC 55900
|||||

QY 292 LeuAaPryrGULyLysIleIySaenGlnThrGlyAlaSerGlyProLySaSerLyLeu 311
|||||
DB 55901 CTGGACTATGAGAAAGATTAAAGAACGATCGAGCATCGGGCCAAAGCAAACTC 55960
|||||

QY 312 LybLeuLhIleLeuGULyLrProAenGUProValProAlaValSerGlnGlyGln 331
|||||

Db 55961 AAGCTCTGCACTGGAGAGCAATGAACTGCTCCCTGCTGTCTCAGAGGGTGGACAG 56020
 Qy 332 TysSerGlnThrProLeuSerProProCysAlaAspSerAlaThiSerGlnAlaAlaGly 351
 Db 56021 AAAAGGAGACCGCCCTCAGTGCACCTGTGCGACTGTGCTCACTCAAGAGCAGACAGA 56080
 Qy 352 GlnArgProValHisProAlaSerValProSerValProSerValGlnProSerLeuLeu 371
 Db 56081 CAAAGGCCCTGATCCCGCCAGCGGCCAGCGTCCAGCGTGCAGCCGTCGCTGTGA 56140
 Qy 372 GluAspSerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGln 391
 Db 56141 GAGGACAGCCCGCTGGTACAGCGCGCTCAGTGGGCTGCACCTGTCCGACAGACAGCGTGGAA 56200
 Qy 392 AspSerAsnLeuLeuValArgSerPheSerLeuAspIleLysSerValSerTyrSerAla 411
 Db 56201 GACAGCAATAGCTCAAGCGGTCCTTCTCTGGATATCAAAATCAGTTTCATATTCAGCC 56260
 Qy 412 SerMetAlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGlnTyrTyrLys 431
 Db 56261 AGCATGGACATCTTACATAGCTCTTCTCTCATCAAGAGATGCTTTGGAAATATCTACAA 56320
 Qy 432 ProSerThrThrLeuAspGlyThrAsnLysLeuGlnPheSerProValGlnGlnLeu 451
 Db 56321 CCTTCACACTCTGGATGGAGCAACAGATATGCGACATTCCTCCCTGTCAGGAACTA 56380
 Qy 452 SerGlnGlnThrProGlnThrSerProAspArgGlnGlnAlaSerIleProLysLeu 471
 Db 56381 TCGAGACACATCTCCCAACCAAGCTCTGTATAGAGAGAACCAAGCATCCCAAGAGCTG 56440
 Qy 472 GlnThrAlaArgProSerAspSerGlnSerLysArgLeuHisSerValArgThrSerSer 491
 Db 56441 CAGACCGCAGGCTTCACAGCAGCAGACAGCAAGCATTCGCTGCAAGACAGCAGC 56500
 Qy 492 SerGlyThrAlaGlnArgSerLeuLeuSerProLeuHisArgSerGlySerValGlnAsp 511
 Db 56501 AGTGGACACCGCCAGAGGTCCTTTATCTTCACATGATCAAGATGGAGCGTGGAGGAC 56560
 Qy 512 AsnTyrHisThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSer 531
 Db 56561 AATTACCAACACAGCTCTCTTTGGGCTTTCCACACAGCAGCAGCAGCAGCTCAGAGCT 56620
 Qy 532 AlaGlyLeuGlyLeuLysGlyTyrPheHisAspIleLeuAlaProGlnThrSerThrPro 551
 Db 56621 GCTGGCCTGGGCTTAAGGGCTGGCACTGGATATCTTGGCCCCCAAGCTCTACCCCT 56680
 Qy 552 SerLeuThrSerSerTyrPheAlaThrGlnSerSerHisPheTyrSerAlaSerAla 571
 Db 56681 TCCCTGACACAGAGCTGATATTTGGCAGAGAGTCTCACACTTCACTGCTCCTCAGCC 56740
 Qy 572 IleTyrGlyGlySerAlaSerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGly 591
 Db 56741 ATCTACGGAGGAGGAGGAGCTTCTGCTGCTACACTGACAGCAGCTGCCACTTGGCGGA 56800
 Qy 592 AspGlnValTyrSerValArgArgGlnLysProSerAspArgAlaAspSerArgArg 611
 Db 56801 GACCAAGTATATCTCTGGCAGCGCGGAGGAGCAAGTACAGAGCTGCTCGGCGCG 56860
 Qy 612 SerTrpHisGlnGlnSerProPheGlnLysGlnPheLysArgArgSerCysGlnMetGln 631
 Db 56861 AGCTGGCATGAAAGAGAGCCCTTTGAAAAAGCAGTTTAAACGAGAAGCTGCCAATATGGA 56920
 Qy 632 PheGlyGlnSerIleMetSerGlnAsnArgSerArgGlnGlnLysValGlySer 651
 Db 56921 TTTGAGAGAGCATATGTCAAGAAAGGTCACGGGAAAGACTCTGGGAAAGTGGGCGAGT 56980
 Qy 652 GlnSerSerPheSerGlySerMetGlnIleIleGlnValSer 665
 Db 56981 CAGTCTAGCTTTTGGGGCAGCATGGAATCATTTAGGCTCTCC 57022
 RESULT 22
 AC131617/c

LOCUS AC131617 188344 bp DNA linear HTG 17-SEP-2002
 DEFINITION Homo sapiens clone RP11-613M22, *** SEQUENCING IN PROGRESS ***, 7
 accession AC131617
 version AC131617.3 GI:22953886
 keywords HTG; HTGS PHASE1; HTGS ACTIVEFIN.
 source Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 188344)
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F., Allen, C.,
 Albrooks, S.L., Amarantunga, H.C., Are, J.R., Ayale, M., Banks, T.,
 Barbarta, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P.,
 Bunay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Dem, A.L., Ding, X., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escoto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
 Homsi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maneswari, M., Mapa, P., Martin, R., Martinale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabhat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokwoko, S., Ogun, M., Okunonu, G.,
 Orsanye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shoochitari, N., Sison, I.,
 Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Syatek, A., Taborp, P., Tamerisa, A., Tameias, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, R., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinrock, G. and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 188344)
 Worley, K.C.
 Direct Submission
 Submitted (25-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 188344)
 Worley, K.C.
 Direct Submission
 Submitted (17-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Sep 17, 2002 this sequence version replaced gi:22538362.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Drafting Center Code: BCM
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information

REMARK	Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
COMMENT	NIH-MGC Project URL: http://mgc.ncl.nih.gov Contact: MGC help desk

Email: c9abbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kwois, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
 Series: IRAK Project: 51 Row: 1 Column: 21
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Source

CDS

```

1..2807
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5176724"
/tissue_type="Brain, lung, Testis, adult, pooled whole"
/clone_lib="NH_MGC_115"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
<1..1852
/codon_start=2
/product="Unknown (protein for IMAGE:5176724)"
/protein_id="AAH1643.1"
/db_xref="GI:21594973"
/translation="CSMLKRLQDDQKVLITELHSAKHKVIDSGKVVVVVQSSD
DVALSDGCVPLVGLKESPVNHLAONGKEGYVLNASTCEKDFIPESTFTEL
LVANIGPTRLPLVLYAGCDVLYNKLKNSGCVLHCLAGISRTALAYIMKRMDEL
VNSDFEKLIPWLWDKVDIEAKNSGCVLHCLAGISRTALAYIMKRMDEL
DEARFQSEKPTLSPICSDKTSAAAGRVHPASVPSVDSVPSLIEPSPLVQALSG
VSEAGVSEKPTLSPICSDKTSAAAGRVHPASVPSVDSVPSLIEPSPLVQALSG
HLSDAPVLEKSLKRSGLDKSVYSAMAASLFGSSSDALEYKPSDRLVQKST
LCGSPVQELSEQPTKPKKEASIPKQLQTPRPSDSKRLSVRTNSGTRQSL
LSPFHRGSEVDNHTSPFLRGISTOOHLTKSAGLGLKMSHDLAPQTSPLTSSW
YFAEASHPYASAKYGGSAASVAYSQGLPTCGQVYVSVRQKPSDRASRSMBE
SSPEKQPKRSKSGMGESIMEINSRELKQVSSQSSFGSMIEIVS"

```

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	2807
Score:	371.00	Matches:	613
Percent Similarity:	99.03%	Conservative:	0
Best Local Similarity:	99.03%	Mismatches:	3
Query Match:	55.75%	Indels:	6
EB:	9	Gaps:	0

US-10-029-345A-109 (1-665) X BC031643 (1-2807)

QY	5	CYSSELYSEULEMELTYSAIRYGRLENGINSINAPLYSVALLEILETHGDLULEMLE	63
Db	2	TGCTCCAACTTATAGACGAAGTTGCACACGACCAAGTGTATATTACAGACTATC	61
QY	70	GLNHISSEALALYSHISLYSVALSPILLEAPCYSSERGLINLYSVALVALVALTYTASP	89
Db	62	CAGCATTCACGCAACATTAAGGTATGACATATGATTACATGCAAGAGTGTGATTTACGAT	122
QY	90	GLNSERSERGLINAPVALASERLEUSESSERASAPCYSPHELETHRVALLEULENGLY	103
Db	122	CAAGCTCCCAAGATATTGCTCTCTCTCTTACAGCTGTTTCTACAGTACTTGGGT	181

QY	110	lyleuenglulyserserthemservaihi sleuleuual aglygilyphenialagluhpser	129
Db	182	AAACTGGAGAAGGCTTCAACTCTGTTCACCTGCTTGAGAGTGGGTTTGCCTGAATTCCT	241
QY	130	ArgCysPhePro ¹ lyleuCysglulglyyserserThleuValPro ¹ hCysIleSerGln	149
Db	242	CGTTGTTTCCCTGGCCTCTGTGAAGAAATCCACTAGTCCCTTACTCGATTCTTCAG	3010
QY	150	ProCysleuProValAlaasnilleglyPro ¹ ThArgilleuP ¹ roasnleuTYrleugly	169
Db	302	CCCTGCTTACTGTTGGCAACATTGGGGCAACCCGAAATCTTCCCAATCTTATCTTGGC	361
QY	170	CysGluArgAspValleuasnlysgileuule-GlnadlnasnGlylleglyTYrValle	189
Db	362	TGCGACGAGATGTCTCTCAACAAGAGCTGAT-GCAGCAGATGGAGATTGGTTATGTGTT	420
QY	189	uaenalaserTYr-ThrCysProlysp ¹ roasphenilleProgluSer ¹ hAspheleuArgy	209
Db	421	AAATGCCAGCAA-TACTGTCCAAAGCCTGACTTTATCCCGAGTCCATTCTCGGTG	479
QY	209	alProValasmpserPheCysglulysilleuPro ¹ ThleuAsp ¹ lyserValAsp	229
Db	480	TGCTGTGATGACAGCTTTGTGTGAATAATTTGCGGTGTGACAAATTCAGTAATT	539
QY	229	heilleglulysalalysalaserasnGlyCysValleuValhCysleualaglyIleS	249
Db	540	TCATTGAGAAAGCAAAAGCTTCATGTGATGTGTTATGTCACGTGTTAGCTGGAGTCT	599
QY	249	etrArgserAlaThrIlealallealyrillemylshgMetAspMcSerleuAspG	269
Db	600	CCGCGTCCGCCACCATCGCTATCGCTCAATCAATGAAGAGATGGAATGCTTTGATG	659
QY	269	IuAlaTYrArgPheVallysglulysArgPro ¹ hIleSerProasnPheAspheleuG	289
Db	660	AACTTACGATTTTGTGAAGAAAAAACACTTACTATATCTCCAACTTCAATTTTCTGG	719
QY	289	IyGlnleuLeuaspTYrGluyslysilleysaenglnThrglyAlaserGlyProlys	309
Db	720	GCCAACTCTGGACTATGAGAAAGAAATTAAAGAACAGCTGGAGCATCAGGGCCAAAG	779
QY	309	erlyserleuysleuenuhisleuGlyusP ¹ roasnGluProValProAlaValSerGlnG	329
Db	780	GCAAACTCAAGCTGCTGCACTCGAGAAAGCCAAATGAACCTGTCCGTCTGTCAAGG	839
QY	329	IyGlyGlnlyserGlnThrP ¹ roleuSerPro ¹ ProCysAlaAspSerAlaThrSerGlu	349
Db	840	GTCGACAGAAAAGCAGAGCGCCCTCAGTCCACCTGTGGCGACTGTCTACTCTCAAGG	899
QY	349	IAlaAglyGlnArgProValhIsp ¹ roAlaserValProserValGlnProS	369
Db	900	CAGCAGAGAAAGGCCCGTCGATCCCGCACCGGCCACGTGCCAGGTCACACCGT	959
QY	369	erleuLeuGlnAspSerProleuValGlnAlaLeuSerGlyLeuuhIleuSerAlaAspA	389
Db	960	CGGTGTGAGAGACAGCCCGCTGTGTACAGGGCTCACTGGCTGTGCACCTGTCCGACAGA	101
QY	389	rgleuGlnAspSerAsnlysleuIysArgSerPheSerleuAspilleyservalSerT	409
Db	1020	GGCTGGAAAGACGCAATAGCTCAAGCGTTCCTCTCTGTGATATCAATTCAGTTTCAT	1071
QY	409	lySerAlaserMetAlaIaSerleuuhisgilyPheSerSerSerGluAspAlaIleuGluT	429
Db	1080	ATTCAGCCAGCATGGCAGCATCTTCACTGAGCTTCTCTCATACAGAGATGCTTGGAAAT	113
QY	429	lyTYrlyAsp ¹ roserThrThrleuAspGlyThrAsnlysleuCysGlnPheSerProValG	449
Db	1140	ACTACAAACTTCCACTACTCTGGATGGAGCAACAAGACTATCGAGTCTCCCGCTTCC	119
QY	449	IngluleuSerGluGlnThrProgluThrSerProAsp ¹ lysgluGluAlaserIleProL	469
Db	1200	AGGAATCTACGAGCAGACTCCGAAACAGTCTCTGATTAAGAGAGCAGAGATCTCCA	1251

QY 469 yelyleuGlnThrAlaArgProSerAspSerGlnSerIysValGluHisSerValArgT 489
 DB 1260 AGAGCTGCAGACCGCAGGCTTTCAGACAGCCAGCAAGCATTCGATCGTCAGAA 1319
 QY 489 hSerser-SerGlyThrAlaGlnArgSerLeuLeuSerProLeuHisArgSerGlySer 508
 DB 1320 CCAGCAAA-CAGTGGCACCAGCCAGAGTCCCTTTTATCTCCATCGATCGAATGGGAGC 1378
 QY 509 ValGluAspAsnThrIleThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeu 528
 DB 1379 GTGGAGACCAATTAACCAACAGCTTCTTTCGCGCTTTCACAGCCAGCAAGCACTTC 1438
 QY 529 ThrIysSerAlaGlyLeuGlyLeuLysGlyTrpHisSerAspIleLeuAlaProGlnThr 548
 DB 1439 AGCAATCTGCTGGCTGGCTGGCTTAAAGGCTGGCACTCGATATCTTGGCCCCCAGACC 1498
 QY 549 SerThrProSerLeuThrSerSerTrpIlePheAlaThrGluSerSerHisPheIysSer 568
 DB 1499 TCTACCCCTTCCCTGCAGCAGCACTGTGATTTTGCACAGAGTCTCACACTTCTACTCT 1558
 QY 569 AlaSerAlaIleThrGlyGlySerAlaSerTrpSerAlaThrSerCysSerGlnLeuPro 588
 DB 1559 GCTTCAGCATCTACAGAGGAGGAGTGCAGTACTCTGCTACAGCTGCAGCCAGCTGCC 1618
 QY 589 ThrCysGlyAspGlnValIleTrpSerValArgArgArgGlnIleProSerAspArgAlaAsp 608
 DB 1619 ACTGGCGAGACCAAGTCTTCTTCTGCGCAGCGGCGAGAGCCAAAGTACAGAGCTGAC 1678
 QY 609 SerArgArgSerTrpHisGluGluSerProPheGlnLysGlnPheIysArgArgSerCys 628
 DB 1679 TCGCGCGAGGCTGCGATGAGAGAGAGCCCTTTGAAAGCAGTTTAAACGACAGAGCTGC 1738
 QY 629 GlnMetGluPheGlyGlyLeuSerIleMetSerGlnAsnArgSerArgGluGluLeuGlyLys 648
 DB 1739 CAATGAAATTGGAGAGAGCATCATGTCAAGAAACAGTCAAGGAGAGCTGGGAGAA 1798
 QY 649 ValGlySerGlnSerSerPheSerGlySerMetGluIleIleGluValSer 665
 DB 1799 GTGGGAGTCACTGCTTTCGCGCAGCATGTGAATCATTTAGGCTTCC 1849

RESULT 24
 AX482372 5111 bp DNA linear PAT 16-AUG-2002
 LOCUS Sequence 41 from Patent WO02057460.
 DEFINITION AX482372
 ACCESSION AX482372
 VERSION AX482372.1 GI:22316945
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 Todderud,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G.,
 Suchard,S., Banas,D., Bassolino,D., Feder,J., Kyslek,S.,
 Mcatee,P., Mintier,G., Siemers,N., Jackson,D.G. and Ramanathan,C.
 TITLE Polynucleotides encoding human phosphatases
 JOURNAL Patent: WO 02057460-A 41 25-JUL-2002;
 SQUIBB BRISTOL MYERS CO (US)
 FEATURES
 source
 1. 5111
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ALIGNMENT SCORES:
 Pred. No.: 0 Length: 5111
 Score: 350.00 Matches: 661
 Percent Similarity: 98.80% Conservative: 0
 Best Local Similarity: 98.80% Mismatches: 4
 Query Match: 52.63% Indels: 8
 DB: 6 Gaps: 0

US-10-029-345A-109 (1-665) x AX482372 (1-5111)
 QY 1 MetAlaHisGluMetIleGlyThrGlnIleValIleThrGluArgValAlaLeuLeuGlu 20
 DB 470 ATGGCCCATGAATGATGATGAACTCAATTTGATGAAAGGTGGTGGCTGCTGGAA 529
 QY 21 SerGlyThrGluLysValLeuLeuIleLeuSerArgProPheValGluTrpAsnThrSer 40
 DB 530 AGTGAACCGGAAAAGTGTGCTAATGATTAACCGGCCATTTGTGAAATACATATCATC 589
 QY 41 HisIleLeuGlnAlaIleAsnIleAsnGlySerLysLeuMetLysArgArgLeuGlnGln 60
 DB 590 CCAATTTGGAAGCCATTAATATCACTCTCCAGCTTATGACGGAAGTGGCAACAG 649
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
 DB 650 GACAAAGTGTAAATTAACAGGCTCATCCAGCATTCAGCCAAACATTAAGTTGACATTTAT 709
 QY 81 CysSerGlnLysValValIleValIleTrpAspGlnSerSerGlnAspValAlaSerLeuSer 100
 DB 710 TGCAGTCAGAGGTTGTATTCATTCATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 769
 QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
 DB 770 GACTGTTTCTCACTGTACTTCTGGGTAACTGAGAAAGCTTCAACTCTGTTCACTG 829
 QY 121 LeuAlaGlyGlyPheAlaLeuPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
 DB 830 CTTCAGGAGGGTTGCTAGGTTCTCTGTTTTCCTGCGCTCTGTAAAGGAAATCC 889
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLysProValAlaAsnIleGlyProThr 160
 DB 890 ACTTAGTCTCTACTGCTGCTTCTAGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 949
 QY 161 ArgIleLeuProAsnLeuTrpLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
 DB 950 CCAATCTTCCCAATCTTATCTTGGCTGCCAGGAGATGTCCTCAACAAGAGCTGAT- 1008
 QY 181 -GlnGlnAsnGlyIleGlyTrpValLeuAsnAlaSerTrp-ThrCysProLysProAsp 200
 DB 1009 GCAGCAGATGGATGGATTGATGTGTTAATGCCAGCAA-TACCTGTCCAAAGCTGACT 1067
 QY 200 IleIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle 220
 DB 1068 TTATCCCCAGTCTCATTTCTCGCGTGCCTGTGAATACACCTTTGTGAAAATTT 1127
 QY 220 euProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAsnGlyCysVal 240
 DB 1128 TCCCGTGTGGACAAATCAGTAGATTTTATTGAGAAACCAAAAGCTCCAAATGATGTG 1187
 QY 240 AlLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTrpIle 260
 DB 1188 TTCTGTGACCTGTTTACGTGGGATCTCCCGTCCGCCACATCGCTATCGCTTACATCA 1247
 QY 260 eLysArgMetAspMetSerLeuAspGluAlaTrpArgPheValLysGluLysArgPro 280
 DB 1248 TGAAGAGATGACATGCTTTTAGATGAAGCTTACAGATTTGTGAAGAAAAGACCTTA 1307
 QY 280 hIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAsp-TyrGluLysLysIleLys 299
 DB 1308 CTATATCTCAAACTTCATTTTCTGGGCCAACTCTGTGC-CTATGAGAAAGATTAA 1366
 QY 300 AsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHis-LeuGluLys 319
 DB 1367 AACCAAGTGGAGCATCAAGGCGCAAGCAAACTCAACTGCTGTC-CTGTGAGAGGC 1425
 QY 319 AsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerPro 339
 DB 1426 AATGAACTGTCTCTGCTGCTTCAAGAGGTGACAGAAAAGCAAGCCCTCTGCTGC 1485
 QY 339 oProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAla 359
 DB 1486 ACCCTGTGCGACTGTGCTACTCAGAGGACAGCAAGAAAGCCGCTGATCCGCCAG 1545

QY 359 rValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAl 379
Db 1546 CGTGGCCAGGCTGCCAGGCTGCAGCCGCTGTTTAAAGAGACAGCCCGTGTACAGGC 1605
QY 379 aLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSe 399
Db 1606 GCTCAGTGGGCTGACCTGTCGCCAGACAGCTGTGAAGACAGCAATTAAGCTCAAGGCTTC 1665
QY 399 rPheSerLeuAspLysLeuSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisG 419
Db 1666 CTCTCTCTGGAATACAAATCAATTCATTAATCAGCAGATGAGACATCTTACATG 1725
QY 419 yPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThreLeuAspGlyL 439
Db 1726 CTCTCTCTATCAAGAAATGCTTGGAAATTAATCAAACTTCCACTACTCTGATGGAGAC 1785
QY 439 rAnLysLeuGlyGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSe 459
Db 1786 CAACAAGCTATGCGCAGTCTCCCTGTTGAGAACTATCGGAGCAGACTCCGAAACGAG 1845
QY 459 rProAspLysGluGluAlaSerLysLeuProLysLysLeuGlnThrAlaArgProSerAspSe 479
Db 1846 TCTGTATTAAGAGAAAGACAGCATCCCAAGAACTGACAGACCGCAGGCTTCAGACAG 1905
QY 479 rGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLe 499
Db 1906 CAGAGACAAAGCATTCATTCGCTCAGAACACAGACAGACTGGCAGCCGCCAAGGCTCCT 1965
QY 499 uLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPh 519
Db 1966 TTTATCTCCACTGCAATCGAAGTGGAGCGTGGAGGCAATTAACACACAGCTTCTCTTT 2025
QY 519 eGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTr 539
Db 2026 CCGCCTTTCACACAGCAGCAGACCTCAAGAACTGCTGCGCTTGAAGGCGTGG 2085
QY 539 pHisSerAspLysLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrTrpH 559
Db 2086 GCACCTCGAATCTTGCGCCGCCACAGCTCTACCCCTTCCAGACAGAGCTGTGATTT 2145
QY 559 eAlaThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTy 579
Db 2146 TCCCAAGAGTCTCTACACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2205
QY 579 rSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgAr 599
Db 2206 CTCTGCTTCAAGCTGACAGCAGCTGCGCACTTGCAGAGACCAAGTCTATTTCTGCGAG 2265
QY 599 gArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSerProPh 619
Db 2266 GCGGCGAAGAGCCAGAGCAGAGCTGACTCGCGGGGAGCTGGCATTGAAGAGCCCTT 2325
QY 619 eGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerTlleMetSerG 639
Db 2326 TGAAGAAGAGCTTTAAACGAGAGCTGCGCAATTTGAATTTGAGAGAGCATCATGTGAGA 2385
QY 639 uAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMe 659
Db 2386 GAAACAGCTACCGGAGAGAGCTGGGAAAGTGGGAGCTGAGTCTTCTTCTGCGGAGAGAT 2445
QY 659 rGluLysLeuGluValSer 665
Db 2446 GGAATTCATTGAGTCTCC 2464

RESULT 25
AX921917
LOCUS AX921917 2071 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 257 from Patent WO02068649.
ACCESSION AX921917
VERSION AX921917.1 GI:40215410
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Patent: WO 02068649-A 257 06-SEP-2002;
Curegen Corporation (US)
FEATURES
source location/Qualifiers
1..2071
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1,94e-288 Length: 2071
Score: 304.00 Matches: 658
Percent Similarity: 98.65% Conservative: 0
Best Local Similarity: 98.65% Mismatches: 7
Query Match: 45.71% Indels: 9
DB: Gaps: 0
US-10-029-345A-109 (1-665) x AX921917 (1-2071)
QY 1 MetaLHisGluMetLysGlyThrGlnLysValThrGluArgLeuValAlaLeuLeuGlu 20
Db 61 ATGGCCCTGATGATGATTTGAACTCAATTTGTTACTGAGAGTTGGCTCTCTGGA 120
QY 21 SerGlyThrGluLysValLeuLeuLysAspSerArgProPheValGluTyrAsnThrSer 40
Db 121 AGTGAACGGAAGAAAGTCTCTAATTGATGATGCGGCGCATTTGGAATCAATATCATCC 180
QY 41 HisLysLeuGluAlaIleAsnLysSerLysLeuMetLysArgArgLeuGln 60
Db 181 CACATTTTGAAGCATTAATATATCACTGCTCCAAAGCTTATGAAGCGAAGGTTGCAACAG 240
QY 61 AspLysValLeuLysThrGluLeuLysHisSerAlaLysHisValAspLys 80
Db 241 GACAAAGTTAATTTACAGAGCTATCCAGCATTCACGAAACATAAGTTGACATTGAT 300
QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 301 TGCAGTCAGAAAGTTGATGATTAAGATCAAAAGCTCCCAAGATGTTGCTCTCTTCA 360
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db 361 GACTGTTTCTCACTGACTTCTGAGTAACTGGAAGAGAGCTTCAACTGTTCACCTG 420
QY 121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuGluGlyLysSer 140
Db 421 CTTCAGAGTGGGTTTGTGAGTTCTCTGTTTTCCTGCTGCTGTAAGGAAATCC 480
QY 141 ThrLeuValProThrCysLysSerGlnProCysLeuProValAlaAsnLysGlyProThr 160
Db 481 ACTCTAGTCTCTCACTGCAATTTCTCAGCTTGTGCTTCACTTGCACCAATTTGGGCCAAC 540
QY 161 ArgLysLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuLysGluLeu 180
Db 541 CGAATTTCTCCAAATCTTATCTTGTGCTGCACAGAAATGCTTCAACAGAGAGCTGAT- 599
QY 181 -GlnGlnAsnGlyLysGlyTyrValLeuAsnAlaSerTyr-ThrCysProLysProAsp 200
Db 600 GACAGCAATGGAGTGGTATGATGATTAATGACAGAA-TACTGTGCCAAACCTGACT 658
QY 200 HeLysProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysLys 220
Db 659 TTAATCCCGAGTCTCAATTTCTCTGCTGCTGCTGATGACACACTTTTGAAGAAATTT 718
QY 220 euProThrLeuAspLysSerValAspPheLysGluLysAlaLysAlaSerAsnGlyCys 240
Db 719 TGCCTGCTTGAACAAATCAAGATTTCTTATGAAGAAAGGCTCCAAATGAGATGTG 778
QY 240 alleuValHisCysLeuAlaGlyLysSerArgSerAlaThrIleAlaIleAlaTyrLleu 260

Db 779 TTCTATGACATCTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCTTACATCA 838
 Qy 260 eTlysaArgMetSerLeuAspGluAlaTyrArgPheVallyrGluYsaArgProt 280
 Db 839 TGAAGAGATGACATGCTTTAGATGAGCTTACAGATTGTTGAAAAAGAACCTTA 898
 Qy 280 hTlSerProAnPheAnPheLeuGlyGlnLeuLeuArgTyrGlyYlyValIleLysA 300
 Db 899 CTATATCTCCAACTTCAATTTCTGGGCCAACTCTGACATATGAGAAAGATTAAGA 958
 Qy 300 sngInThrGlyAlaSerGlyProLysSerLysLeuLysLeuLysLeuGlyLysPro 320
 Db 959 ACCAGACTGGAGATCAGGGCCAAAGACAACTCAAGCTGCTGACCTGGAGAACCCA 1018
 Qy 320 sngLubProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeuSerPro 340
 Db 1019 ATGAACTGTCCTGCTGCTCAGAGGGTGGACAGAAAGAGAGAGCCCTCAGTCCAG 1078
 Qy 340 rOcYalAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerY 360
 Db 1079 CCTGTCCGACTGTGCTACTCAGAGGACAGACAGACAAAGCCCTGTCATC--C--C--G 1132
 Qy 360 aLProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAla 380
 Db 1133 --CC--AGCTAACCCAGGCTGACAGCCCTGCTTTAGAGACAGCCCGCTGTGACAGGCC 1189
 Qy 380 euSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAnlyLeuLysArgSerP 400
 Db 1190 TCAGTGGGCTGACACTGTCCGACAGACAGGCTGGAGACAGCATTAAGCTCAAGCTTCC 1249
 Qy 400 heSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyP 420
 Db 1250 TCTCTGTGATATCAATCAGTTCAATTCATTCACCCAGACAGACAGATCTTTCATGCT 1309
 Qy 420 heSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThr 440
 Db 1310 TCTCTCATCAAGAGATGCTTTGGAATCTACAAACCTTCACTACTGTGATGGAGACA 1369
 Qy 440 sNlyLeuCyGlnPheSerProValGlnLubLeuSerGluGlnThrProGluThrSerP 460
 Db 1370 ACAAGCATACCAAGTCTCCCTCGTTGAGAACTATCGAGACAGACTCCCGAANACAGTC 1429
 Qy 460 rOAspLysGluGluAlaSerIleProLysLeuGlnThrAlaArgProSerAspSerG 480
 Db 1430 CTGATTAAGAGGAGGACAGCATCCCAAGAGCTGCAGACGCCGAGCTTTCAGACAGCC 1489
 Qy 480 InSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 Db 1490 AGAGCAAGCGATTCGATTCGATCAAGACAGACAGTGGACCCGACAGAGGTCCCTTT 1549
 Qy 500 euSerProLeuHisArgSerGlySerValGluAspAnlyrHisSerPheLeuPheG 520
 Db 1550 TATCTCCACGATCGAAGTGGAGGAGCTGGAGACAAATTACACACAGCTCTCTTTG 1609
 Qy 520 LysLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLysLeuGlyLysGlyTyrP 540
 Db 1610 GCTTTTCCACACACGACAGACACTCAGAAAGTCTGCTGCGGCTTAAAGGGCTGCG 1669
 Qy 540 IAspSerAlaIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheA 560
 Db 1670 ACTGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGAGCTGGATTTTG 1729
 Qy 560 lATHrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrS 580
 Db 1730 CCAAGAGTCTCACTTCTTACTGCTCCACAGCATCTAGAGAGGAGGAGGAGGAGTTACT 1789
 Qy 580 eAlaTyrSerCySerGlnLeuProThrCySerGlyAspGlnValTyrSerValArgArgA 600
 Db 1790 CTCCTACAGCTGACAGCTGCGCCACTTGGGAGACCAAGTCTATTCGTGCGCAGCG 1849
 Qy 600 rGlnLysProSerAspArgAlaAspSerArgArgSerTyrPheGluGluSerProPheG 620

Db 1850 GGCAGAACCAAGTAGACAGACTGACTCGCGCGAGAGCTGGCATGAAGAGAGCCCTTTG 1909
 Qy 620 lUlyGlnPheLysArgArgSerCyGlnMerGluPheGlyGluSerIleMetSerGluA 640
 Db 1910 AAAAGCATTTTAAACCAAGACTGCAATGGAATTTGAGAGACATCATGTACAGAGA 1969
 Qy 640 sNAspSerArgGluGluLeuGlyLysValGlySerGlnSerPheSerGlySerMetG 660
 Db 1970 ACAGTCAAGGAGAGAGCTGGGAGAGTGGGAGTGCATGATGACTTTTGGGACAGATG 2029
 Qy 660 lUllelGluValSer 665
 Db 2030 AAATCATTTAGGCTCTCC 2046
 Db 2030 AAATCATTTAGGCTCTCC 2046
 RESULT 26
 BC042101
 LOCUS
 DEFINITION
 Homo sapiens dual specificity phosphatase 16, mRNA (CDNA clone
 MGC:50665 IMAGE:440399), complete cde.
 ACCESSION
 BC042101 GI:27469788
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 3284)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buell, K.H., Scheffer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heich, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Caaveant, T.L.,
 Schectz, T.E., Brownstein, M.J., Uebli, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A.C., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalsky, U., Small, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 JOURNAL
 MEDLINE
 PUBMED
 22389257
 12477932
 2 (bases 1 to 3284)
 Strausberg, R.
 Direct Submission
 Submitted (23-DEC-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-ehg.stanford.edu>
 Contact: (Dickson, Mark) mdickpax1.stanford.edu
 (Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAX Plate: 88 Row: a Column: 17

QY 440 snlyseuCyseGlnPheSerProValGlnIuleuSerGlnGlnInThrProGluThrSerP 460
DB 1641 ACAAGCTATGCCAGTCTCCCTGTTGAGAACTATCGAGAGCACTCCGGAACCAAGTC 1700
QY 460 roApLyseGlnIuValaSerIleProLyseLysLeuGlnThrAlaArgProSerAspSerG 480
DB 1701 CTGATAAGGAGGAAGCAGCATCCCAAGAAAGCTGACAGCCGACGGCTTTCAGACAGCC 1760
QY 480 InseLyseArGluHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuL 500
DB 1761 AGAGCAAGCATGTCATTCGTCAGAACCCAGACAGCTGACCGCCCAAGAGTCCCTTT 1820
QY 500 euseProLeuHlsArgSerGlySerValGlnuAspAntYrHlsPheSerPheLeuPheG 520
DB 1821 TATCTCCACTGCTCATCGAAGTGGAGCGTGGAGGACATTTCCACACCAAGCTTCTTTTCG 1880
QY 520 lYleuSerThSerGlnGlnHisLeuThrLysSerValaGlyLeuGlyLeuGlyTyrH 540
DB 1881 GCCTTTCACACAGCAGACACACTCAAGAGTCTGCGCTGGGCTTAAGGGCTGGC 1940
QY 540 lseSerPheLeuAlaProGlnInThrSerThProSerLeuThrSerSerTyrPheA 560
DB 1941 ACTCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGGATTTTG 2000
QY 560 lathrgIuseSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrS 580
DB 2001 CCAAGAGTCTCACTTACTCTGCTCAGCACTTACGAGGAGGAGGAGGAGGAGGAGTACT 2060
QY 580 exrAlaTyrSerCyseSerGlnLeuProThrCyseGlyAspGlnValTyrSerValArgAr 600
DB 2061 CTGCTTACAGCTCAGCAGCAGCTCCCACTTGGGAGACCAAGCTATTCTGTCGGCAGGC 2120
QY 600 rglGlnLysPProSerAspArgAlaAspSerArgArgSerTyrHlsGlnLysPProPheG 620
DB 2121 GGCAGAGAGCCAGATGACAGAGCTGACTCGCGCGAGCTGGCATGAAGAGCCCTTTTG 2180
QY 620 lYleuGlnPheLysArgArgSerCyseGlnMetGlnPheGlyGlySerIleMetSerGln 640
DB 2181 AAAAGCATTTAAAGCAGAAAGCTGCCAAATGGAATTTGAGAGAGCATATATGACAGA 2240
QY 640 snArgSerArgGlnIuLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetG 660
DB 2241 ACAAGTCAAGGAGAGAGCTGGGGAAGTGGGAGTCACTTCTTTGGGAGAGCATG 2300
QY 660 lYleGlnIuValSer 665
DB 2301 AAATCATTTGAGTCTCC 2317
RESULT 27
AX835305 1916 bp DNA linear PAT 15-DEC-2003
LOCUS Sequence 2429 from Patent EP1347046.
AX835305
VERSION AX835305.1 GI:39921440
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Iseogi,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yamanoto,J.T., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahara,K. and
Masuno,Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 2429 24-SEP-2003;
Research Association for Biotechnology (JP)
FEATURES
Source
1..1916
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Alignment Scores:
Pred. No.: 1,42e-164 Length: 1916
Score: 178.00 Matches: 278
Percent Similarity: 99.29% Conservative: 0
Best Local Similarity: 99.29% Mismatches: 1
Query Match: 26.77% Indels: 2
DB: 6 Gaps: 0
US-10-029-345A-109 (1-665) x AX835305 (1-1916)
QY 387 AlaAspArgLeuGlnuAspSerAnlyseLysLysArgSerPheSerLeuAspIleLysSer 406
DB 142 GCAGAGAGCTGGAGACAGCAATTAAGCTCAAGCGTCTCTCTCTGATATCAATCA 201
QY 407 ValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPheSerSerSerGlyuAspAla 426
DB 202 GTTTCATATTACCGCCAGCATGGCAGCATCTTAACAATGCGCTTCTCATCAGAAAGATGCT 261
QY 427 LeuGlnTyrTyrLysProSerThrThrLeuAspGlyThrAsnLysLeuCyseGlnPheSer 446
DB 262 TTGGAATTAATAAAGCTTTCATCTGATGGAGCAACAGCTATGCGAGTCTCC 321
QY 447 ProValGlnIuleuSerGlnGlnInThrProGluThrSerProAspLyseGlnuAlaSer 466
DB 322 CTGTTCAAGAACTATCGAGAGCACTCCGAAACAGTCTGATTAAGAGAGAGAGCCAGC 381
QY 467 lIeProLyseLysLeuGlnInThrAlaArgProSerAspSerGlnSerLysArgLeuHisSer 486
DB 382 ATCCCAAGAGAGTGGAGACCGCAGCGCTTTCAGACAGCAGCAGCAGCAGCAGTTCG 441
QY 487 ValArgThrSerSerSerGlyThrAlaGlnArgSerLeuLeuSerProLeuHisArgSer 506
DB 442 GTCCAGAACAGACAGCAGTGGCAACCGCCAGAGGTCCCTTTTATCTCAGTCAAGAT 501
QY 507 GlySerValGlnuAspAntYrHisThrSerPheLeuPheGlyLeuSerThr-SerGlnG 526
DB 502 GGGAGGTGGAGAGCAATTAACACACAGCTTCTTTGGCTTTCAT-CAGCCAGCA 560
QY 526 nHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrHlsAspAspIleLeuAlaP 546
DB 561 GCACTCAAGAGTGTGCTGGCTGGCTTAAAGGCTGGCACTCGATATCTTGGCCCC 620
QY 546 GlnInThrSerThrProSerLeuThrSerSerTyrPheAlaThrGlnSerSerHisP 566
DB 621 CCAGACCTTACCCCTTCCCTGACCAAGCAGCTGTATTTTGCACAGAGTCTCACACTT 680
QY 566 eTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSerAlaTyrSerCyseSerG 586
DB 681 CTACTCTGCTCAGCCATCTACGAGAGCAGTGCACATTACTTGCCTTACAGCTGCAGCCA 740
QY 586 nLeuProThrCyseGlyAspGlnValTyrSerValArgArgArgGlnLysProSerAsp 606
DB 741 GCTGCCCATTTGGGAGAGCAAGTCAATTTCTGTGGCAGGCGGCAAGGCAAGTACAG 800
QY 606 galaAspSerArgArgSerTyrHlsGlnLysPProPheGlnLysPheLysArgArg 626
DB 801 AGCTGACTCGCGCGGAGCTGGCATGACAGAGCCCTTTGAAACAGATTAAAGCAG 860
QY 626 gSerCyseGlnMetGlnPheGlyGlySerIleMetSerGlnuAspArgSerArgGlnL 646
DB 861 AAGCTGCCAAATGGAATTTGAGAGGACATCATGTAGAGAACAGGCTACGGGAAAGAGCT 920
QY 646 uGlyLysValaGlySerGlnSerSerPheSerGlySerMetGlnIleGlnIuValSer 665
DB 921 GGGGAAGTGGGAGTCACTTACGTTTTCGGGACAGCATGAATCATTTGAGTCTCC 978
RESULT 28
AK098310 1916 bp mRNA linear PRI 15-JUL-2002
LOCUS AK098310
DEFINITION Homo sapiens cDNA FLJ40991 fis, clone UTRU015198, weakly similar
to DUAL SPECIFICITY PROTEIN PHOSPHATASE 8 (BC 3.1.3.48).
ACCESSION AK098310

VERSION AK098310.1 GI:21758299
 KEYWORDS oligo capping; flis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagahtsuna,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshina,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuno,Y., Nagai,K. and Isogai,T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1916)
 AUTHORS Isogai,T. and Yamamoto,J.
 JOURNAL Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
 FEATURES
 source
 1. 1916
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UTERU2015198"
 /issue_type="uterus"
 /clone_id="UTERU2"
 /note="Cloning vector: pME18SFL3"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,42e-164 Length: 1916
 Score: 178.00 Matches: 278
 Percent Similarity: 99.29% Conservative: 0
 Best Local Similarity: 99.29% Mismatches: 1
 Query Match: 26.77% Indels: 2
 DB: 9 Gaps: 0
 US-10-029-345A-109 (1-665) x AK098310 (1-1916)
 QY 387 AlaAspArgLeuGluAspSerAsnLysLeuLysArgSerPheSerLeuAspIleLeuLysSer 406
 DB 142 GCGAGCAGCTCGAAGACAGCAATTAAGCTTCAAGCGTTCTCTCTGGAATTCAAATCA 201
 QY 407 ValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPheSerSerSerGluAspAla 426
 DB 202 GTTTCATATTACGCCGATGCGATGCGATCTTACATGCTTCTCTCATCAGAAATGCT 261
 QY 421 LeuGluTyrTyrTyrProSerThrThrLeuAspGlyThrAsnLysLeuGlnPheSer 446
 DB 263 TTGGAAATCTACAAACCTTCCACTCTGATGGAGCAACAAACCTTATGACAGTTCTCC 321
 QY 447 ProValGlnGluLeuSerGluGlnThrProGluTyrThrProAspLysGluGluAlaSer 466
 DB 322 CCTGTTCAGAACTATCGAGAGACACTCCCGAAACCAAGTCTGATTAAGAGAGAGACCCAGC 381
 QY 467 IleProLysLysLeuGlnThrAlaArgProSerAspSerGlnSerLysArgLeuHisSer 486
 DB 382 ATCCCAAGAAAGCTCAGACCGCCAGGCTTTCAGACAGCCAGACAGAGATGATTCG 441
 QY 487 ValArgThrSerSerSerGlyThrAlaGlnArgSerLeuSerProLeuHisArgSer 506

DB 442 GTCAGAACAGCAGCAGTGGACCGCCAGAGATCCCTTTATTCCTCATCGATCGAAGT 501
 QY 507 GlySerValGluAspAsnTyrHisThrSerPheLeuPheGlyLeuSerThr-SerGlnG1 526
 DB 502 GGGAGCGTGGAGAGCAATTAACACACAGCTTCTTTCCGCTTCCAT-CAGCCAGCA 560
 QY 526 HisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheSerSerAspIleLeuAlaPr 546
 DB 561 GCACCTCAGAAAGTCTGCTGGCTGGCTGGCTTAAAGGCTGGACCTCGATATTTGGCCCC 620
 QY 546 GlnThrSerThrProSerLeuThrThrSerSerTyrTyrPheAlaThrGlnSerSerHisAsp 566
 DB 621 CCAGACCTCTACCCCTTCCCTGACAGCAGCGCTGATTTGGCAGACAGATCCTCACACTT 680
 QY 566 eTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSerAlaTyrSerCysSerG1 586
 DB 681 CTACTCTGCTCAGCACTACCTGAGGAGGAGGAGCCAGTACTCTGCTCAGCTGAGCA 740
 QY 586 nLeuProThrCysGlyAspGlnValTyrSerValArgArgGlnLysProSerAspAr 606
 DB 741 GCTGCCCACTTCGCGAGACCAAGTATTTCTGTGGCAGCGCGAGCAAGCCAGAGCAG 800
 QY 606 gAlaAspSerArgArgSerTyrPheGlyGlnGluSerProPheGlnLysGlnPheLysArgAr 626
 DB 801 AGCTGACTCGGGCGGAGCTGGCATGAGAGAGCCCTTTGAAAGCAGTTTAAACGCAG 860
 QY 626 gSerGysGlnMetGluPheGlyGlnSerTlleMetSerGluAsnArgSerArgGlnGluLe 646
 DB 861 AAGCTGCCAAATGGAATTTGAGAGAGCATATGTCAGAGAAACAGTCCAGGAAAGACT 920
 QY 646 uGlyLysValGlySerGlnSerSerPheSerGlySerMetGlnLleLeuGluValSer 665
 DB 921 GGGGAAATGGGAGCTAGTACTGTTTCCGGCAGCATGAGAAATCATTAGCTCTCC 978
 RESULT 29
 AC097818 244605 bp DNA linear HTG 10-MAY-2003
 LOCUS Rattus norvegicus clone CH230-52K22, WORKING DRAFT SEQUENCE.
 AC097818
 DEFINITION AC097818.7 GI:30520485
 ACCESSION AC097818
 VERSION
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 244605)
 AUTHORS Muzny,D.,Marie, Metzker,M.,Lee, Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Albrooks,S., Amin,A., Angiano,D., Anyalbechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Bunay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Centner,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dedrich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Dudin,K., Dvali,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Goady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamli,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M., Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpachy,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,U., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Louised,H., Lozad,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,

Manqum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Mijavajevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Nundasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwakoelameh, O., Okunolu, G., Olarunbasogun, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
 Plapper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, R., Ruz, S.,
 Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartbeyn, A., Sison, I., Sitter, C.D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J.,
 Stead, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanat, K.,
 Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willison, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G., and Gibbs, R.A.

Unpublished
 Direct Submission
 2 (bases 1 to 244605)
 Worley, K.C.

Direct Submission
 Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 244605)
 Rat Genome Sequencing Consortium.
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:24819653.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Project name: GFRH
 Center project name: GFRH
 Center clone name: CH230-52K22

----- Summary Statistics
 Assembly program: Atlas 3.0:
 Consensus quality: 234323 bases at least Q40
 Consensus quality: 235890 bases at least Q30
 Consensus quality: 236794 bases at least Q20
 Estimated insert size: 241565; sum-of-contigs estimation
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have

* provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 1 244605: contig of 244605 bp in length.
 Location/Qualifiers
 source
 1. 244605
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-52K22"
 complement(165..989)
 /note="clone boundary
 clone end: Sp6
 site: EcoRI
 end sequence: BH316904"
 23405..124172
 /note="clone boundary
 clone end: T7
 site: EcoRI
 end sequence: BH316901"
 238775..239794
 /note="wgs_end_extension
 clone end: T7"
 243232..244605
 /note="wgs_end_extension
 clone_end:T7"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,2e-57 Length: 244605
 Score: 71.00 Matches: 71
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 10.68% Indels: 0
 DB: 2 Gaps: 0

US-10-029-345a-109 (1-665) x AC097818 (1-244605)
 QY 7 GlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGluGluSerGlyThrGluVal 26
 Db 123608 GGAACCAATATGTTATCTAGAGGTTGTGGCTCTGTGGAAAGCGCACCGAAAGTG 123667

QY 27 LeuLeuIleASPserArgProPheValGluTyrAsnThrSerHisIleLeuGluAlaIle 46
 Db 123668 CTGCTAATATGACAGCGGCGCATTGTGGAGTACAAATACGTCACATTCTGGAAAGCATC 123727

QY 47 AsnIleAsnCySerIysLeuMetIysARGArgLeuGlnIleAspIysValLeuIleThr 66
 Db 123728 AATATCACTGCTCCAAACTGATGAGCGAAGGCTCCAAACAAAGATTATTAACA 123787

QY 67 GluLeuIleGlnHisSerAlaIysHisIysVal 77
 Db 123788 GAACATCATCCAGCATTCGCAAAACATAAGGTA 123820

RESULT 30
 AB052157
 LOCUS
 DEFINITION Mus musculus MKP-7 mRNA for MAP kinase phosphatase-7, complete cds.
 AB052157
 ACCESSION
 VERSION AB052157.1 GI:13990988
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 Masuda, K., Shima, H., Watanabe, M. and Kikuchi, K.
 MKP-7, a novel mitogen-activated protein kinase phosphatase,
 functions as a shuttle protein
 J. Biol. Chem. 276 (42), 39002-39011 (2001)
 JOURNAL MEDLINE 21486429
 PUBMED 11489891

REFERENCE
2
AUTHORS Masuda, K.
TITLE Comparative chromosomal mapping of MKP-7 genes
JOURNAL Unpublished
3 (bases 1 to 4874)
AUTHORS Masuda, K., Shima, H. and Kikuchi, K.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2000) Kouhei Masuda, Institute for Genetic
Medicine, Division of Biochemical Oncology and Immunology, Kita-ku
Kita5-1-10, Nishi-7-kyoumei, Sapporo, Hokkaido 060-0815, Japan
E-mail: khouei@imm.hokudai.ac.jp, Tel: 81-11-706-5536,
Fax: 81-11-707-6839

FEATURES

source
1. 4874
/organism="Mus musculus"
/mol_type="mrna"
/db_xref="taxon:10090"
/tissue_type="cerebellum"
1. 4874
/gene="MKP-7"
433. 2415
/gene="MKP-7"
/codon_start=1
/product="Map kinase phosphatase-7"
/protein_id="BAB47240.1"
/translation="MAHEMIGIOIVTESI,VALLESGTEKVLIDSRPFVNTSHILE
/db_xref="GI:13990989"
/translation="MAHEMIGIOIVTESI,VALLESGTEKVLIDSRPFVNTSHILE
AININSGKIMKRLQODKYLITELIIONSAKHKVIDIDCNQKVVYVQSSDVGSLSDC
FITVLGKLERFNSVHLIAGAEFSPGFCGKSTLVPTCLISQPLPVANIGPT
RLPLNLYICQDRVDLNKMQNGIGYVRCNANSTCEKPTPIESHLFVANDSCFK
ILPWLKSDYDFLEKASNGCVLIHCLAGISRSATIAIAYIMKRMDSIDEAYRVEKE
KRPITSPNFMGQMDYKTIINOTNGSGPKSKLHLIDKPEFVPAPASGGMKSA
LSLSPCANSTSEASGORLVHPASVSPLOSLIPSLVQALSGIOLSEKLESDTKL
KRSFIDIKSVSYASMASTLGFSESELDJCKRSATITDGNKICQSPVQVEYEQS
PETSPPKEXEANI PKOPPPRPSQVTLRLSVRTSSGSGTQRPFSPLRSGVEDNY
HNPFLGSLTSQQLRTKSAIGLKGWHSIDILAPQSSAPSLTSSWTFATEPSLVSASA
IYGGNSYSAVSCGLPTCSQDIYSVRROKPTDRASRRSRHESPEFKQPKRRSCQ
MEFSGSIMSENSREIKGVGSSPSGSMETIEVS"

ORIGIN

Alignment Scores:
Pred. No.: 4,22e-55 Length: 4874
Score: 67.00 Matches: 266
Percent Similarity: 95.00% Conservative: 0
Best Local Similarity: 95.00% Mismatches: 7
Query Match: 10.08% Indels: 14
DB: 10 Gaps: 0

US-10-029-345A-109 (1-665) x AB052157 (1-4874)

QY 15 LeuValAlaLeuLeuGluSerGlyThrGluValLeuLeuIleAspSerAlaProPhe 34
Db 475 TTGGTGGCTCTCTCTGGAAGGTGAACGGAAGAAAGTGCTGCTAATTGATAGCGACCATTT 534
QY 35 ValGluTyrAsnThrSerHisIleLeuGluAlaIleAsnIleAsnCyseSerIysLeuMet 54
Db 535 GTGGAATACAAATACGCTCACAATTTTGGAAAGCATTAATATCACTGCTCCAAACTGATG 594
QY 55 LysAlaGlyGluGlnGlnAspIysValLeuIleThrGluLeuIleGlnHisSerAlaIys 74
Db 595 AAGCAAGGTTGCCAACGACAAAGATTAATTAACAACTTATCCACATTCCTCCAAAG 654
QY 75 HisLysValAspIleAspCyseSerGlnLysValValValTyrAspGlnSerSerGlnAs 94
Db 655 CATTAAGTTGACATTGACTGCAA-TCAGAAAGGTGGAGTTATGATCAAAAGTTCCCAAGA 713
QY 94 pValAla-SerLeuSerSerAspCysePheLeuThrValLeuLeuGluIysLeuGluIys- 113
Db 714 TGTCTGG-TTCTCTGTCGTACAGACTGCTTTCTCAGCTGTAATCTTGGGTAACGAGAG-G 771
QY 114 SerPheAsnSerValHisLeuLeuAlaGlyIlePheAlaGluPheSerArgCysePhePro 133
Db 772 AGCTTCAACTCTGTCCACCTGCTTGCAGGTGCTTGGAGTGTCTCTGTTGTTCCCT 831

QY 134 GlyLeuCyseGluGlyIysSerThrLeuValProThrCyseIleSerGlnProCyseLeuPro 153
Db 832 GGCTCTGTGAAGGAAGCACTGACCTAGCTCCTACCTGCATATCTCAGCTTGCTTACT 891
QY 154 ValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyrLeuGlyCyseGlnArgAsp 173
Db 892 GTTGGAACTTGGGCCCAACTGATTTCTTCCACTCTCTATCTTGCGTCCAGGAGAT 951
QY 174 ValLeuAsnLysGlu-LeuIle-GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerT 193
Db 952 GTCTTCAACAAAGA-CCTGAT-GCAACAGAAATGGGATTTGGTATGTTAATGACAGCA 1009
QY 193 Yr-ThrCyseProLysProAspPheIleProGluSerHisPheLeuArgValProValAsn 212
Db 1010 A-TTACCTGTCCAAAGCTGACCTTCACTACCTGATCTCTCTGCGAGTCTGTGAT 1068
QY 213 AspSerPheCyseGluLysIleLeuProThrPheAspIysSerValAspPheIleGluIys 232
Db 1069 GACAGCTTTTGTGAGAAATCCTACCATGTTGGACAAAGTCTGTGATTTTACGAAA 1128
QY 233 AlaLysAlaSerAsnGlyCyseVal-LeuValHisCyseLeuAlaGlyIleSerArgSerAl 252
Db 1129 GCMAAAGCTTCAATGCTGTGCTTATC-CACTGCTTACGCTGATCTCTGCTCCGC 1187
QY 252 AThrIleAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArg 272
Db 1188 CACTATTCCTATTGCTTCACTATCATATGAGAGATGACATGCTCTAGATGAGGCTTACAG 1247
QY 272 GPeValIysGluIysArgProThrIleSerProAsnPheAsnPhe 287
Db 1248 ATTGTGAAGAAAAAAGACCTACATATCTCCGAATTTTATTTT 1293
RESULT 31
BC059232
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 4943)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Dietzenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schnee, T.E., Brownstein, M.J., Udell, T.B., Tothyluk, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunartne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smailus, D.E.,
Scherer, A., Schein, J.B., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
MEDLINE
PUBMED
12477932
2 (bases 1 to 4943)
Strausberg, R.
REFERENCE
TITLE
JOURNAL
Submitted (01-OCT-2003) National Institutes of Health, Mammalian

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schlier, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schmeitz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, Y., Hu, Y., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.T., Skalska, U., Small, D.E., Schmeitz, A., Schein, J.E., Jones, S.J., and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 4975)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (02-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Stevan Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susana Chan, Readman Chiu, Chris Fjell, Erin Garlund, Ran Guin, Leticia Hsiao, Martin Krzywinski, Keta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice Maceady, Steven Ness, Pawan Pandoh, Anna-Liisa Pirahnu, Parvaneh Saeedi, Jacqueline Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRAX Plate: 126 Row: n Column: 2
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 16485489.

location/Qualifiers

1..4975

organism="Mus musculus"

molecule="mRNA"

strain="C57BL/6"

db_xref="taxon:10090"

clone="MGC:66822 IMAGE:6830600"

tissue_type="Brain, enriched mouse brain 12.5dp"

clone_lib="NIH-BMAP_F00"

lab_host="DH10B"

note="Vector: pYX-ASC"

1..4975

gene="Dusp16"

note="synonyms: MKP-7, MKP7, D6Ert213e"

db_xref="locusID:70686"

db_xref="MGI:1917936"

565..2547

codon_start=1

product="Dusp16 protein"

protein_id="AAH57321.1"

CDS

/db_xref="GI:3498087"

/db_xref="locusID:70686"

/translation="MAHMIIGTQITVLSVALLSEGTKEVLLIDSRPEVYNTSHILE
AIVNICKLKRRLQDDKVLITQLQSHAKRKVIDCNQKVVYDQSSQDVSSSDC
FLTVLGLKERSFNSVHLIAGFAFSGFPGICGKSTLPTCTSCPLPVANIGPT
RLTPVLGSCQDVLDMOONGGVYVANSNTCPKDFIPESHFLPVNDSCCK
ILPMDKSVDFEKXAKNGCVILKLAGISRAITLAIYIMKMDMLDAVPRVKE
KRPTSPNPNMGOLMDYEKTIINNOTGSGKSTIKLHIDKPSRPVAAEGRKSA
LSLSPCANSTSEAGQVLHPVAVPSIQPSLDESPVLQVLSGLQSSKEBDSKL
KRSFSLDIKSVSYSAKMASLHGSSEFALDYKPSATLIDCTNKLQPSPVQESQ
PETSDDKEAHIPKQPPRPSQVTSVHVRKSSGSLTPRFPSPHRSQVSDNY
HTNPLFGISTQOHLTKSAGLKGMSHDIAPQSSAFLSSWYFATPEPSHLVSA
ITGNSYSVAVSCGQLPTPCSPQIVSVRRORPTPADSRBSWHESPREKQPKRSCQ
MEFGSIENSEKRSREILKVSQSFSFSMTIIVS"

613..963

note="RHOD; Region: Rhodanese Homology Domain"

/db_xref="CDD:smat0450"

1036..1452

/note="RSPc; Region: Dual specificity phosphatase,
catalytic domain. Ser/Thr and Tyr protein phosphatase.
The enzyme's tertiary fold is highly similar to that of
lysine-specific phosphatases, except for a 'recognition'
region"

/db_xref="CDD:pfam00782"

ORIGIN

Alignment Scores:

Pred. No.: 4,296-55 Length: 4975

Score: 67.00 Matches: 266

Percent Similarity: 95.00% Conservative: 0

Best Local Similarity: 95.00% Mismatches: 7

Query Match: 10.08% Indels: 14

DB: Gaps: 0

US-10-029-345A-109 (1-665) x BC057321 (1-4975)

QY 15 LeuValAlaLeuLeuGluSerGlyThrGluValLeuLeuIleAspSerArgProPhe 34

DB 607 TTGGTGGCTCTGCTGGAAAGGGAACGGAAGAAAGTCTGCTAATGATAGCCGACATTT 666

QY 35 ValGluTyrAsnThrSerHisIleLeuGluAlaIleAsnIleAsnIleAsnIleMet 54

DB 667 GTGGAATCAATCAATGCTCAATTTGGAAGCCATTAATCACTGCTCCAACTGATG 726

QY 55 LysArgArgLeuGlnGlnAspIleValLeuIleThrGluLeuIleGlnHisSerAlaLys 74

DB 727 AAGCGAAGTTGCAACGAGCAAGATTAATTAACAACCTTAATCCAACTTCTGCAAG 786

QY 75 HisLysValAspIleAspCysSerGlnLysValAlaValIleTyrAspGlnSerSerGlnAs 94

DB 787 CATTAAGTTGACATTAATGCTGCAATTTGGAAGCCATTAATCACTGCTCCAACTGATG 845

QY 94 PValAla-SerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGluLys- 113

DB 846 TGTGG-TTCTCTGCTCTGCAAGCTTCTCTCACTGATCTTGGGTAACCTGGAGAG-G 903

QY 114 SerPheAsnSerValHisLeuLeuAlaGlyGlyPheAlaGluPheSerArgCysPhePro 133

DB 904 AGCTTCACTCTGTCACCTGCTGCAAGGAGGCTTGTAGATTTCTCTGTTTCTCT 963

QY 134 GlyLeuCysGluGlyLysSerThrLeuValProThrCysIleSerGlnProCysLeuPro 153

DB 964 GGCCTCTGTGAAGGAAGTCACTCACTGATCCCTACCTGCAATATCTAGCTTGTACT 1023

QY 154 ValAlaAsnIleGlyProThrArgGlyLeuLeuProAsnLeuTyrLeuGlyCysGlnArgAsp 173

DB 1024 GTTGGCAACATTGGGCACTGCAATCTTCCAACTCTATCTTGTGCTGGCAGAGAT 1083

QY 174 ValLeuAsnLysGlu-LeuIle-GlnGlnAsnGlyIleGlyTyrValIleAsnAlaSerT 193

DB 1084 GTCTCAACAAGCA-CTTGAT-GCAACGATAGGATGGCTATGTTAAAGCCAGCA 1141

QY 193 Yr-ThrCysProLysProAspPheIleProGluSerHisPheLeuArgValProValAsn 212


```

Db      1142 A-TACGCTCAAGAGCTGACTTCATACGTAATCTCCTCGAGAGCTGTGAAT 1200
Qy      213 AAGSPPheCysAGLulysIleLeuPOTRleuApslySerylAlaSpPheIleGluYs 232
Db      1201 GACAGCTTTTGTAGAAAATCCACATGCTGTGACAAAGCTGTGATTTTCATTGGAGAA 1260
Qy      233 AAlaAlaSerangIysVal-LeuValHisCysLeuAlaGlyIleSerArgSerI 252
Db      1261 GCAAAAGCTTCATGAGCTGTGCTTATC-CAGCTGTGAGTGTGAGATCTCTCGCTCCG 1319
Qy      252 aThrlleAlaIleAlaTyrlleMetIysArgMetAspMetSerLeuApsIuAlaTyAr 272
Db      1320 CACTATTGCTATTGCTCATCATGAAAGAGATGACATCTCTAGATAGGCTTACAG 1379
Qy      272 gPheValIysGluIysArgProThrlleSerProaPheAsnpe 287
Db      1380 ATTGTGAAGAAAAGAAAGACTTACTATATCTCCGAATTTTAATTTT 1425

RESULT 33
AC133722 242590 bp DNA linear HTG 20-NOV-2002
LOCUS Rattus norvegicus clone CH230-240110, *** SEQUENCING IN PROGRESS
DEFINITION
AC133722
AC133722.2 GI:25138890
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 242590)
Muzny,D.,Marle, Metzker,M.,Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,B., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Byrant,N., Bunay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Hayes,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flaggy,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgie,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunsatne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoguez,M.,
Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Koyar,C.,
Kowalski,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewka,L., Loulsegad,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Miloasavljeyic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munkidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwachukwemeh,O., Okwuonu,G., Olarunpungoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankuch,C.,
Plomper,F., Polindexter,A., Popovic,D., Primus,B., Pu,L.-L.,
Puzo,M., Quirroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soosa,J.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KClE
Center clone name: CH230-240110
----- Summary Statistics
Assembly program: Phrap; Version 0.990329
Consensus quality: 220426 bases at least Q40
Consensus quality: 220448 bases at least Q30
Consensus quality: 223410 bases at least Q20
Estimated insert size: 224964; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
240024: contig of 240024 bp in length
240025: gap of unknown length
240125: contig of 1049 bp in length
241174: gap of unknown length
241274: gap of unknown length
241274: contig of 1317 bp in length.
Location/Qualifiers
1. 242590
/mol_type="Rattus norvegicus"
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-240110"

```

```

misc_feature      1..1094
                  /note="wgs end_extension
                  clone_end:Sp6"
misc_feature      7392..8314
                  /note="clone boundary
                  clone_end:Sp6
                  site:
end sequence:BZ106067"
misc_feature      complement(235121..235967)
                  /note="clone boundary
                  clone_end:T7
                  site:
end sequence:BZ106066"
misc_feature      236448..237947
                  /note="wgs end_extension
                  clone_end:T7"
misc_feature      237998..240024
                  /note="wgs end_extension
                  clone_end:T7"

ORIGIN
Alignment Scores:
Pred. No.:      9,83e-53      Length:      242590
Score:          66.00         Matches:      66
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    9.92%        Indels:      0
DB:             2            Gaps:        0

US-10-029-345A-109 (1-665) x AC133722 (1-242590)

QY      595  TYSerValArArGrArgInLysPProSerAraPAlaSPSerArGrSerTrpHis 614
DB      12283 TATTCTGTTCTGRRGGCGGCAAGCTTAGTGACAGCTGACTCACGGCGGAGCTGGCAAT 12342
QY      615  GIUGLuserProPheGluLysGlnPheLysArGrSerCysGlnMetGluPheGlu 634
DB      12343 GAAAGAGAGCCCTTTGAGAAGCAGTTTAACCGAGAGAGCTGCCAATATGCAATTTGGAGAG 12402
QY      635  SerIleMetSerGluAsnArGrSerArGrGluLeuGlyLysValGlySerGlnSerSer 654
DB      12403 AACATTATATGTGAGAAACAGGTCTGAGGAAAGACTGGGCAAGTGGGAGCCCAATCTTAC 12462
QY      655  PreSerGlySerMetGlu 660
DB      12463 TTTTCAGAGACATGGAG 12480

RESULT 34
AC128093/c.      250782 bp      DNA      linear      HTG 20-NOV-2002
LOCUS            Rattus norvegicus clone CH230-84N7, *** SEQUENCING IN PROGRESS ***.
ACCESSION        AC128093
VERSION          AC128093.3 GI:25137709
KEYWORDS         HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
SOURCE           Rattus norvegicus (Norway rat)
ORGANISM         Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 250782)
Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguitano,D.,
Anyaleechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardens,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Devila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

```

```

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregregis,E., Geer,K., Gill,R., Grady,M., Guerra,T., Guevara,M.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Huylk,S., Hune,J., Idlibird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Muntada,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackemeah,O., Okwoum,G., Olarnunagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,
Plapper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,M., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wiczzyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,U., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Unpublished
2 (bases 1 to 250782)
Worley,K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 250782)
Baylor Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:23265205.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/r2/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GYAC
Center clone name: CH230-84N7
----- Summary Statistics
Assembly program: Phrap; version 0.990329

```


DEFINITION Mus musculus map kinase phosphatase-M B1 isoform mRNA, complete cds.
ACCESSION AF345953
VERSION AF345953.1 GI:13625396
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 4026)
AUTHORS Matsuguchi,T., Musikacharoen,T., Johnson,T.R., Kraft,A.S. and Yoshikai,Y.
TITLE A Novel MKP is an Important Negative Regulator of LPS-mediated JNK Activation in Macrophages
JOURNAL Mol. Cell. Biol. 20, 6999-7009 (2001)
REFERENCE 2 (bases 1 to 4026)
AUTHORS Matsuguchi,T., Musikacharoen,T. and Yoshikai,Y.
TITLE Direct Submission
JOURNAL Submitted (07-FEB-2001) Medicine, Nagoya University, 65 Tsurumai-cho, Nagoya, Aichi 466-8550, Japan
FEATURES
source location/Qualifiers
1..4026
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/cell_type="macrophage"
558..1625
/note="MKP-M B1 isoform; protein phosphatase"
/codon_start=1
/product="map kinase phosphatase-M B1 isoform"
/protein_id="AAK35054.1"
/db_xref="GI:13625397"
/translation="MAHEMIGQIVTESIVALLSEGTREKULLIDSRPFVNTSHILEAININCKIMKRRLOQDRVLITELIQSAKHVIDDNQKYVVDSSQDVGSLSDCFLIVLGLDEKSFNSVHLAGEFSEFCPEGLCEGSTLVPTCTISQCLPVANIGPTRIPLNLYLGCQDVNLKMDQONGIGVLANSTCPKPDPIPESHFLRPVNDSCFKILPMDISVDFIEKAKASNGCVLIHCLAGISRSATIAIAYIMKMDMLDEAYRROKLTBRADSRRTGMRAPLKSISLNAEAKMNLERALCRRTPGRSMARMAASPAAMRSRSLEKTSLLLTFLVFPVHK"

ORIGIN
Alignment Scores:
Pred. No.: 3,21e-52 Length: 4026
Score: 64.00 Matches: 64
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.62% Indels: 0
DB: 10 Gaps: 0
US-10-029-345A-109 (1-665) x AF345953 (1-4026)

QY 114 SerPheAsnSerValHisIleuLeuAlaGlyIleuPheAlaGluPheSerArgCysPhePro 133
DB 897 AGCTTCAATCTGTCTCCACCTGCTGAGGAGGCTTTCGAACTCTCTGTTGTTCCCT 956
QY 134 GlyLeuCysGluGlyIlySerSerThiLeuValProThrCysIleSerGlnProCysLeuPro 153
DB 957 GGCCTCTGGAAGAAAGTCCACTAGTCCTACTGATATCTCAGCGTTGCTTAACT 1016
QY 154 ValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAsp 173
DB 1017 GTTGGAAACATTGGGCAACTGAAATCTTCCCAATCTCTATCTTGCGTCAGCGAAT 1076
QY 174 ValLeuAsnIly 177
DB 1077 GTCTCAACACAG 1088

RESULT 37 AF345951 4992 bp mRNA linear ROD 24-SEP-2001
LOCUS AF345951
DEFINITION Mus musculus map kinase phosphatase-M A1 isoform mRNA, complete cds.

ACCESSION AF345951
VERSION AF345951.1 GI:13625392
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 4992)
AUTHORS Matsuguchi,T., Musikacharoen,T., Johnson,T.R., Kraft,A.S. and Yoshikai,Y.
TITLE A Novel MKP is an Important Negative Regulator of LPS-mediated JNK Activation in Macrophages
JOURNAL Mol. Cell. Biol. 20, 6999-7009 (2001)
REFERENCE 2 (bases 1 to 4992)
AUTHORS Matsuguchi,T., Musikacharoen,T. and Yoshikai,Y.
TITLE Direct Submission
JOURNAL Submitted (07-FEB-2001) Medicine, Nagoya University, 65 Tsurumai-cho, Nagoya, Aichi 466-8550, Japan
FEATURES
source location/Qualifiers
1..4992
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/cell_type="macrophage"
558..2591
/note="MKP-M A1 isoform; protein phosphatase"
/codon_start=1
/product="map kinase phosphatase-M A1 isoform"
/protein_id="AAK35052.1"
/db_xref="GI:13625393"
/translation="MAHEMIGQIVTESIVALLSEGTREKULLIDSRPFVNTSHILEAININCKIMKRRLOQDRVLITELIQSAKHVIDDNQKYVVDSSQDVGSLSDCFLIVLGLDEKSFNSVHLAGEFSEFCPEGLCEGSTLVPTCTISQCLPVANIGPTRIPLNLYLGCQDVNLKMDQONGIGVLANSTCPKPDPIPESHFLRPVNDSCFKILPMDISVDFIEKAKASNGCVLIHCLAGISRSATIAIAYIMKMDMLDEAYRROKLTBRADSRRTGMRAPLKSISLNAEAKMNLERALCRRTPGRSMARMAASPAAMRSRSLEKTSLLLTFLVFPVHK"

ORIGIN
Alignment Scores:
Pred. No.: 3,83e-52 Length: 4992
Score: 64.00 Matches: 93
Percent Similarity: 97.89% Conservative: 0
Best Local Similarity: 97.89% Mismatches: 1
Query Match: 9.62% Indels: 2
DB: 10 Gaps: 0
US-10-029-345A-109 (1-665) x AF345951 (1-4992)

QY 194 ThrCysProIySPoAspPheIleProGluSerHisPheLeuArgValProValAsnAsp 213
DB 1137 ACCGTGCAAGCTGACTTCACTCAATCTCACTTCTGAGAGGCTGGAATGAC 1196
QY 214 SerPheCysGluIySileuProTyrPleuAspIlySerSerValAspPheIleGluYsAla 233
DB 1197 AGCTTTGTGAGAAATCTCAACATGAGTGGAGCAAGCTGTGATTTCAATGAAAGCA 1256
QY 234 LysAlaSerAsnGlyCysVal-IeuValHisCysLeuAlaGlyIlySerArgSerAla 253
DB 1257 AAAGCCTCCAAATGCGTGTGCTTATC-CACCTCTTAGCGGAGATCTTCGCTCGCAC 1315
QY 253 rIleAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPh 273
DB 1316 TATTGCTATTGCTCACTCATGAAAGAGATGACATGCTCTTGAGATGAGGCTTACAGATT 1375
QY 273 eValIyGluIyArgProThrIleSerProAsnPheAsnPhe 287

Db 1376 TGTGAAGAAAAAGACTATATCTCCGAATTTATTTT 1418

RESULT 38

AC126692

LOCUS 201474 bp DNA linear HTG 12-OCT-2002

DEFINITION Mus musculus chromosome UNK clone RP23-246J18, WORKING DRAFT

AC126692

SEQUENCE, 9 unordered pieces.

AC126692

AC126692.2 GI:23915608

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 201474)

JOURNAL McPherson, J.D. and Waterston, R.H.

REFERENCE

AUTHORS The sequence of Mus musculus clone

TITLE Unpublished

JOURNAL 2 (bases 1 to 201474)

REFERENCE

AUTHORS McPherson, J.D. and Waterston, R.H.

TITLE Direct Submission

JOURNAL Submitted (08-JUL-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE

AUTHORS 3 (bases 1 to 201474)

TITLE McPherson, J.D. and Waterston, R.H.

JOURNAL Direct Submission

Submitted (12-OCT-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

COMMENT

On Oct 12, 2002 this sequence version replaced gi:21703684.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: submissions@wustl.wustl.edu

Project Information -----

Center project name: M.BA0246J18

----- Summary Statistics -----

Sequencing vector: M13; 0%

Sequencing vector: plasmid; 100%

Chemistry: Dye-terminator ET; 0% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 197971 bases at least Q40

Consensus quality: 198628 bases at least Q30

Consensus quality: 199034 bases at least Q20

Insert size: 195000; agarose-fp

Insert size: 200674; sum-of-contigs

Quality coverage: 9.67 in Q20 bases; agarose-fp

Quality coverage: 9.99 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. -----

1 2368: contig of 2368 bp in length

2369 2468: gap of unknown length

2469 6036: gap of 3568 bp in length

6037 6136: gap of unknown length

6137 14367: contig of 8231 bp in length

14368 14467: gap of unknown length

14468 28520: contig of 14053 bp in length

28521 28621: gap of unknown length

28622 41185: contig of 12555 bp in length

41186 41286: gap of unknown length

41287 57620: contig of 16335 bp in length

57621 57721: gap of unknown length

57722 74576: contig of 16856 bp in length

FEATURES

source

1. 201474

Location/Qualifiers

1. 201474

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="UNK"

/clone="RP23-246J18"

1. 2368

/note="assembly_name:Contig14"

2469. 6036

/note="assembly_name:Contig15"

6137. 14367

/note="assembly_name:Contig16"

14468. 28520

/note="assembly_name:Contig17"

28621. 41185

/note="assembly_name:Contig18"

41286. 57620

/note="assembly_name:Contig19"

57721. 74576

/note="assembly_name:Contig20"

74677. 125817

/note="assembly_name:Contig21"

125918. 201474

/note="assembly_name:Contig22"

ALIGNMENT Scores:

Pred. No.: 7.53e-50 Length: 201474

Score: 63.00 Matches: 63

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 9.47% Indels: 0

DB: 2 Gaps: 0

US-10-029-345A-109 (1-665) x AC126692 (1-201474)

QY 15 LeuValAlaLeuLeuGluSerGlyThrGluValLeuLeuLeuAspSerArgProPhe 34

Db 66762 TTGGTGGCTCTGCTGGAAAGTGAACGAAAAGTCTCTATATTATAGCCACCATTT 66821

QY 35 ValGluTyrAsnThrSerHisIleLeuGluAlaIleAsnIleAsnCysSerIysLeuMet 54

Db 66822 GTGGAATACAAATACGCTCACATTTTGGAAAGCCATTATATCAACTGCTCAAACTGATG 66881

QY 55 LysArgArgLeuGlnGlnIleAspIysValLeuIleThrGluLeuIleGlnHisSerAlaIys 74

Db 66882 AAGCGAAGGTTGCACAGACAAAGATATTATTCAGAACTTATCCAACTTCTGCAAG 66941

QY 75 HisIysVal 77

Db 66942 CATTAAGTA 66950

RESULT 39

AC126692/c

LOCUS 201474 bp DNA linear HTG 12-OCT-2002

DEFINITION Mus musculus chromosome UNK clone RP23-246J18, WORKING DRAFT

AC126692

SEQUENCE, 9 unordered pieces.

AC126692

AC126692.2 GI:23915608

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 201474)

JOURNAL McPherson, J.D. and Waterston, R.H.

REFERENCE

AUTHORS The sequence of Mus musculus clone

TITLE Unpublished

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (08-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 201474)
McPherson,J.D. and Waterston,R.H.
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (12-OCT-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Oct 12, 2002 this sequence version replaced gi:21703684.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M BA0246J18

----- Summary Statistics -----
Sequencing vector: MJ3, 0%
Sequencing vector: plasmid, 100%
Chemistry: Dye-terminator Big Dye, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 197971 bases at least Q40
Consensus quality: 198628 bases at least Q30
Consensus quality: 199034 bases at least Q20
Insert size: 195000; agarose-fp
Insert size: 200674; sum-of-contigs
Quality coverage: 9.62 in Q20 bases; agarose-fp
Quality coverage: 9.99 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2368: contig of 2368 bp in length
* 2369 2468: gap of unknown length
* 2469 6036: contig of 3568 bp in length
* 6037 6136: gap of unknown length
* 6137 14367: contig of 8231 bp in length
* 14368 14467: gap of unknown length
* 14468 28520: contig of 14053 bp in length
* 28521 28620: gap of unknown length
* 28621 41185: contig of 12565 bp in length
* 41186 41285: gap of unknown length
* 41286 57620: contig of 16335 bp in length
* 57621 57720: gap of unknown length
* 57721 74576: contig of 16856 bp in length
* 74577 74676: gap of unknown length
* 74677 125817: contig of 51141 bp in length
* 125818 125918: gap of unknown length
* 125918 201474: contig of 75557 bp in length.
Location/Qualifiers
1. .201474

FEATURES
source
1. .201474
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="10M"
/clone="RP23-246J18"
misc_feature
1. .2368
/note="assembly_name:Contig14"
misc_feature
2469. .6036
/note="assembly_name:Contig15"
misc_feature
6137. .14367
/note="assembly_name:Contig16"
misc_feature
14468. .28520

misc_feature
28621. .41185
/note="assembly_name:Contig18"
misc_feature
41286. .57620
/note="assembly_name:Contig19"
misc_feature
57721. .74576
/note="assembly_name:Contig20"
misc_feature
74677. .125817
/note="assembly_name:Contig21"
misc_feature
125918. .201474
/note="assembly_name:Contig22"
ORIGIN

Alignment Scores:

Pred. No.: 6,97e-48 Length: 201474
Score: 61.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.17% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-109 (1-665) x AC126692 (1-201474)

QY 605 AaPArGAlaAsPSeArGArGSeTrPhisGluGluSeRPropGluVysGluPheLys 624
|||
DB 165111 GACAGAGCTGACTCGAGCGAGCTGCATGAAAGAGAGCCCTTTGAAAAGCAGTTTAA 165052

QY 625 ArgArGSeRyGgImeGluPheGlyGluSerIleMetSerGluAsnArGSeArGlu 644
|||
DB 165051 CGCAGAAAGCTCCCAATGGAATTGGAGAGAGACCTTATGTCGAGAACAGGTCCAGGGAG 164992

QY 645 GluIeuGlyVysValGlySerGlnSerSerPheSerGlySerMetGluIleGluVal 664
|||
DB 164991 GAGCTGGGCAAGGAGGCGAGCAGTTCCTTCGCGACGACGATGATCATGAGGTC 164932

QY 665 Ser 665
|||
DB 164931 TCT 164929

RESULT 40

AC118035

LOCUS

DEFINITION

AC118035

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

CDS

/cell_type="macrophage"
558..1460
/note="MKP-M B2 isoform; protein phosphatase"
/codon_start=1
/product="map kinase phosphatase-M B2 isoform"
/protein_id="AAK35055.1"
/db_xref="GI:13625399"
/translation="MAHEMIGQIVTESVALLSEGTKEVLLIDSRPEVYNTSHILE
AINNCSKLMKRRLQODRVLTLELOHSAKHVLDIDNOKXVVYVQSSODVGSLSDDC
FLTVLLGKLEKRSFNSVHLADLMQONGIGVYVNAATCRKPPPIPSHPLRVVNDSE
CEKILPWLDKSDVLEKAYASNGCVLHCLAGISRSATIAIYIMKRMMSLDEAYRR
OKPTDADSRRTGMRAPLKSILNAEAAKMNLERALCRRTGPRGSWARMWASPAPBAA
MRSRLSEKTSILLTLVLPVHK"

ORIGIN

Alignment Scores:

Pred. No.: 2.03e-41 Length: 3861
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.97% Indels: 0
DB: 10 Gaps: 0

US-10-029-345A-109 (1-665) x AF345954 (1-3861)

QY 613 TPDHAGLUGLuserProphEgIuLySGIuPhElyARgSerCyAGIuMetGluPhe 632

DB 1247 TGGCATGAAGAGAGCCCTTTGAAAAGCGAGTTTAAACGAGAGCTGCCAAATGGAAATT 1306

QY 633 GLYGJuserIleWetSerGluAnaGSerArGJugIuLeuGJLYyVaIglYserGin 652

DB 1307 GGGAGAGCATTAATGTCGAGAAACAGCTCCAGGAGAGAGCTGGGCAAGTGGGAGCCAG 1366

QY 653 SerSerPheSerGlySerMetGluIleIleGluValSer 665

DB 1367 TCCAGCTCTCCGCGAGCATGAGATCATCGAGTCTCT 1405

RESULT 42 4827 bp mRNA linear ROD 24-SEP-2001

AF345952 LOCUS Mus musculus map kinase phosphatase-M A2 isoform mRNA, complete

DEFINITION AF345952

ACCESSION AF345952

VERSION AF345952.1 GI:13625394

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 4827) Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

AUTHORS Matsuguchi,T., Musikacharoen,T., Johnson,T.R., Kraft,A.S. and Yoshikai,Y.

TITLE A Novel MKP is an Important Negative Regulator of LPS-mediated JNK

JOURNAL Activation in Macrophages

AUTHORS Mol. Cell. Biol. 20, 6999-7009 (2001)

JOURNAL Matsuguchi,T., Musikacharoen,T. and Yoshikai,Y.

JOURNAL Direct Submission

COMMENT Submitted (07-FEB-2001) Medicine, Nagoya University, 65

LOCATION Tsunuma-cho, Nagoya, Aichi 466-8550, Japan

FEATURES source

1..4827

/organism="Mus musculus"

/mol_type="mRNA"

/strain="BALB/c"

/db_xref="taxon:10090"

/cell_type="macrophage"

558..2426

/note="MKP-M A2 isoform; protein phosphatase"

/codon_start=1

/product="map kinase phosphatase-M A2 isoform"

/protein_id="AAK35053.1"

/db_xref="GI:13625395"

CDS

/translation="MAHEMIGQIVTESVALLSEGTKEVLLIDSRPEVYNTSHILE
AINNCSKLMKRRLQODRVLTLELOHSAKHVLDIDNOKXVVYVQSSODVGSLSDDC
FLTVLLGKLEKRSFNSVHLADLMQONGIGVYVNAATCRKPPPIPSHPLRVVNDSE
CEKILPWLDKSDVLEKAYASNGCVLHCLAGISRSATIAIYIMKRMMSLDEAYRR
OKPTDADSRRTGMRAPLKSILNAEAAKMNLERALCRRTGPRGSWARMWASPAPBAA
MRSRLSEKTSILLTLVLPVHK"

ORIGIN

Alignment Scores:

Pred. No.: 2.44e-41 Length: 4827
Score: 53.00 Matches: 93
Percent Similarity: 97.89% Conservative: 0
Best Local Similarity: 97.89% Mismatches: 1
Query Match: 7.97% Indels: 2
DB: 10 Gaps: 0

US-10-029-345A-109 (1-665) x AF345952 (1-4827)

QY 194 ThrCysProLysProAspPheIleProGluSerHisPheLeuArgValProValAsnAsp 213

DB 972 AACTGTCCAAAGCCGACTTATACCTGAACTCACTTCCTCGGAGATCCCTGTGATGAC 1031

QY 214 SerPheCysGluValIleLeuProTfLeuAspLysSerValAspPheIleGluLysAla 233

DB 1032 AGCTTTGTGAGAAATCTTCCATGTTGAGCAAGCTGTGAGATTTTCATTGAGACAAAGCA 1091

QY 224 LysAlaSerArgGlyCysVal-LeuValHisCysLeuAlaGlyIleSerArgSerAlaTh 253

DB 1092 AAAGCTTCATAGGCTGTGTCTTATC-CACCTGCTTACCTGGAGATCTCCCTCCGCAC 1150

QY 253 rIleAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgph 273

DB 1151 TATGCTATGTGCTTACATCATGAGAGAGATGACATGCTCTCAATGAGGCTTACAGATT 1210

QY 273 eVallysgIuLysArgProthIleSerProAspPheAsnPh 287

DB 1211 TGTGAAGAAAAAGACCTACTATATCTCCGAATTTTAATTTT 1253

RESULT 43 5047 bp mRNA linear PRI 17-JUN-2003

HS804384 LOCUS Homo sapiens mRNA, cDNA DKFZp451D0719 (from clone DKFZp451D0719).

ACCESSION AL833073

VERSION AL833073.2 GI:30268359

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 5047) Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Bloecher,H., Boecker,M., Mewes,H.W., Well,B., Amd,C., Osanger,A.,

TITLE Direct Submission

JOURNAL Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764

COMMENT On Apr 30, 2003 this sequence version replaced gi:21733664.

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de.

BranscheWig/Germany) within the cDNA sequencing consortium of the

This clone (DKFZp451D0719) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcententrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clonezrp@de.further

information about the clone and the sequencing project is available

at <http://mips.gsf.de/proj/cDNA/>.

LOCATION Ingolstaedter Landstr.1, D-85764

FEATURES source

1..5047

/organism="Homo sapiens"

/mol_type="mRNA"

/strain="Homo sapiens"

/db_xref="taxon:9606"

/cell_type="macrophage"

558..2426

/note="MKP-M A2 isoform; protein phosphatase"

/codon_start=1

/product="map kinase phosphatase-M A2 isoform"

/protein_id="AAK35053.1"

/db_xref="GI:13625395"


```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp451D0719"
/db_xref="taxon:9606"
/map="12p13.2"
/clone="DKFZp451D0719"
/issue_type="human spinal cord"
/clone_idb="451 (synonym: hlccl). Vector psportl; host
DH10B; sites NotI + SalI"
/dev_stages="adult"
5027

ORIGIN
polya_site

Alignment Scores:
Pred. No.: 2,01e-35 length: 5047
Score: 47.00 Matches: 47
Percent Similarity: 100.00% Conservative: 0
Beat Local Similarity: 100.00% Mismatches: 0
Query Match: 7.07% Indels: 0
DB: 9 Gaps: 0

US-10-029-345A-109 (1-665) x HSM804384 (1-5047)

QY 77 ValAspIleApCyseSerGlnYsValIvalYtrAspGlnserSerGlnAspValAla 96
821 GTTGACATTGATTCGACGTCAGAGGTTTGTTACATCAAGCTCCCAAGATGTTGCC 880
DB 821 GTTGACATTGATTCGACGTCAGAGGTTTGTTACATCAAGCTCCCAAGATGTTGCC 880
QY 97 SerLeuSerSerApCysePheLeuThrValIleuLeuGlyYlyLeuLeuGlnYsSerPheAsn 116
881 TCTCTCTCTTCACAGCTGTTTCTCAGCTGCTGGGTAACTGAGAGAGAGAGCTTCAAC 940
DB 881 TCTCTCTCTTCACAGCTGTTTCTCAGCTGCTGGGTAACTGAGAGAGAGAGCTTCAAC 940
QY 117 SerValIleuLeuAlaGly 123
941 TCTGTTCACTGCTGAGGT 961
DB 941 TCTGTTCACTGCTGAGGT 961

RESULT 44
AC097776 247148 bp DNA linear HTG 13-NOV-2002
LOCUS Rattus norvegicus clone CH230-94G4, *** SEQUENCING IN PROGRESS ***
DEFINITION
2 unordered pieces.
AC097776.5 GI:24942830
VERSION HTG: HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 247148)
REFERENCE
1 Muzny, D., Matie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alibrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blych, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
Chacko, U., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dedrich, D.,
Delgado, O., Dengson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgiev, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
Gunnarsson, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladik, S. L., Hodgson, A., Hogue, M.,
Hollins, B., Howell, S., Hu, Y. S., Hume, J., Idelbr, D., Jackson, A.,
Jackson, B., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowle, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

```

```

Lorensbawa, L., Louised, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenan, B.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundaas, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nuckelmeier, O., Okunou, G., Olarnunagsoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
Plopper, F., Ponder, A., Popovic, D., Primus, E., Pu, L., L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smay, D.,
Sneed, A., Sodergren, E., Song, X., Z., Sorrelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K.,
Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, J., Wei, X., White, F.,
Williams, G., Willson, R., Wleczek, R., Wood, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, U., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G., and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 247148)
Worley, K. C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 247148)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:23269395.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GFPZ
Center clone name: CH230-94G4
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 208726 bases at least Q40
Consensus quality: 212396 bases at least Q30
Consensus quality: 214499 bases at least Q20
Estimated insert size: 210868; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
----- NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/GenBank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is

```

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 11276: contig of 11276 bp in length
 * 11277 11376: gap of unknown length
 * 11377 247148: contig of 235772 bp in length.

FEATURES

source
 1.247148
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-94G4"
 1. 1340
 misc_feature
 /note="wgs contig"
 107191..108362
 /note="wgs contig"
 132629..136209
 /note="wgs contig"
 136341..137406
 misc_feature
 /note="wgs contig"

ORIGIN

Alignment Scores:
 Pred. No.: 2,22e-14 Length: 247148
 Score: 27.00 Matches: 27
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.06% Indels: 0
 DB: 2 Gaps: 0

US-10-029-345A-109 (1-665) x AC097776 (1-247148)

QY 97 SerLeuSerSerAspGlyPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheLeu 116
 |||||
 Db 1077 TCTCTGTCCTCAGACGCTTCTCTCCTGCTTGGGTAACTGGAGAGAGCTTCAAC 1136
 |||||
 QY 117 SerValHisLeuLeuAlaGly 123
 |||||
 Db 1137 TCTGTTACCTGCTTGCAGGT 1157
 |||||

RESULT 45
 AK116903 3540 bp mRNA linear INV 30-NOV-2002
 LOCUS Ciona intestinalis cDNA, clone:cieg018b07, full insert sequence.
 ACCESSION AK116903.1 GI:23589623
 VERSION AK116903.1
 KEYWORDS FLI CDNA.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Clona.

REFERENCE
 AUTHORS
 1 Satou,Y., Yamada,L., Mochizuki,Y., Takatori,N., Kawashima,T.,
 Sasaki,A., Hamaguchi,M., Awazu,S., Yagi,K., Sasaki,Y.,
 Nakayama,A., Ishikawa,H., Inaba,K. and Satoh,N.
 A cDNA resource from the basal chordate Ciona intestinalis
 JOURNAL Genesis 33 (4), 153-154 (2002)
 MEDLINE 22191024
 PUBMED 12203911

TITLE 2 (bases 1 to 3540)
 JOURNAL Direct Submission
 REFERENCE Satou,Y. and Satoh,N.
 AUTHORS Submitted (04-OCT-2002) Nori Satoh, Kyoto University, Department of
 TITLE Zoology, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 JOURNAL (E-mail:satou@ascidian.zool.kyoto-u.ac.jp, Tel.81-75-753-4095,
 Fax:81-75-705-1113)
 Ciona intestinalis cDNA Project (URL:
 http://ghost.zool.kyoto-u.ac.jp/indexr1.html).
 Location/Qualifiers

COMMENT
 FEATURES
 source
 1. 3540
 /organism="Ciona intestinalis"

/mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="cieg018b07"

Alignment Scores:
 Pred. No.: 5.73e-11 Length: 3540
 Score: 22.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.31% Indels: 0
 DB: 3 Gaps: 0

US-10-029-345A-109 (1-665) x AK116903 (1-3540)

QY 270 AlaTyrArgPheValysGlyLysArgProThrIleSerProAsnPheAsnPheLeuGly 289
 |||||
 Db 842 GCTTACAGGTTCTGGAAGAAAAAGACCAACCAATTTCTCCCACTTAAATTTCTCGGG 901
 |||||
 QY 290 GlnLeu 291
 |||||
 Db 902 CAGCTG 907
 |||||

RESULT 46
 BC045110 2264 bp mRNA linear PRI 07-OCT-2003
 LOCUS Homo sapiens dual specificity phosphatase 8, mRNA (cDNA clone
 DEFINITION MGC:52239 IMAGE:5478677), complete cds.
 ACCESSION BC045110
 VERSION BC045110.1 GI:28277227
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 1 (bases 1 to 2264)
 Strauberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Bueltow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heideh,F.,
 Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Schaeetz,T.E., Brownstein,M.J., Usdin,T.B., Tohiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loggellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.U., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Wotley,K.C., Hale,S., Garcia,A.M., Gay,L.U., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
 Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932

TITLE 2 (bases 1 to 2264)
 JOURNAL Direct Submission
 REFERENCE Strauberg,R.
 AUTHORS Submitted (27-JAN-2003) National Institutes of Health, Mammalian
 TITLE Gene Collection (MGC), Cancer Genomics Office, National Cancer
 JOURNAL Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT
 NIH-MGC Project URL: http://mgc.ncl.nih.gov
 Contact: MGC help desk
 Email: gcgaps-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILUNI)
 DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigrl.nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dierich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kong, P., Latic, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maki, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantipop, S., Thomas, P.J., Touchman, J.W.,
Tsugeon, C., Vogt, J.D., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
Series: IRAL Plate: 44 Row: 1 Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4758211.

FEATURES

source

1..2264
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:52239 IMAGE:5478677"
/issue_type="Brain, astrocytoma, grade IV"
/clone_lib="NIH MGC_98"
/lab_host="DH10B"
/note="Vector: POTB7"
1..2264
/gene="DUSP8"
/note="synonyms: HB5, HVH-5, HVH8"
/db_xref="locusid:1850"
/db_xref="WMT:602038"
2..1879
/codon_start=1
/product="dual specificity phosphatase 8"
/protein_id="AAH45110.1"
/db_xref="GI:28277228"
/db_xref="locusid:1850"
/translation="MAGDRLPRKVDMAKKLASILRGPGGPIVDSRFSVEYNMHWL
SSVNTCCSKLVKRLQGGVITIAELIQPARSVEVETEDQDVVVYQSTRDASVLAAD
SFSLISLKDGFDSVALITGGFATPSSCPGLCEKPAALLPMSLSQPCIPVPSVG
LRIILPHLYLGSQKVDLMTONGISIVLANASNCSPDPICSEMRPVPINDVYC
EKLLPWLDSIERIDAKLSGCCVYVHCAGISRSRTTIAITMTKMGSSDDAIRFV
KORRSISPNFNLGQLLTERSLKULALQGPSPGTPPEPPSPAAAGALPLPLP
PTSESATGNAARREGSLAGEPAPPPPTPTSAQQLRGHLSSDRLQDTNRLK
SFSLDIKSAVVASRRPDGPPDPGPAKLCIDSPGALGSSPSPDPAPEAR
PPRRRRPPRPPAGPARSPHSLGILNGDAROPRHGLSALSPGLPGGPPAGPAM
APPLDSRGPSPDPGPPGSPGAGAGGVLPAFGAGAGPGGSGDILRRRAAREP
RDARTGWPEEPAPETQFKRRSCOMEPEBGMVEGRANGELALGKQASFGSVETIV
S"

gene

CDS

47..403
/note="RHOD; Rhodanese Homology Domain"
/db_xref="CDD:smar00450"
479..895
/note="DSPC; Region: Dual specificity phosphatase,
catalytic domain. Ser/Thr and Tyr protein phosphatases.
The enzyme's tertiary fold is highly similar to that of
tyrosine-specific phosphatases, except for a 'recognition'
region"
/db_xref="CDD:pfam00782"

ORIGIN

Alignment Scores:

Pred. No.: 3..69e-09 Length: 2264
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.01% Indels: 0
DB: 9 Gaps: 0

US-10-029-345a-109 (1-665) x BC045110 (1-2264)

Qy 242 VALHSCYLeuAlaGlyIleSerArgSerAlaTrilleAlaIleAlaTyrIleMetLys 261
|||||
Db 731 GTCACATGCTGCTGGCATCTCCGCTCTGCACCATCGCATCGCCATCATGAG 790

RESULT 47 AR255944 2377 bp DNA linear PAT 20-DEC-2002
LOCUS AR255944 Sequence 3 from patent US 6482644.
DEFINITION AR255944
ACCESSION AR255944 GI:27305203
KEYWORDS AR255944.1
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2377)
AUTHORS Cowart, L.M.
TITLE Antisense modulation of dual specific phosphatase 8 expression
JOURNAL Patent: US 6482644-A 3 19-NOV-2002;
FEATURES Location/Qualifiers
source 1..2377
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 3.84e-09 Length: 2377
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.01% Indels: 0
DB: 6 Gaps: 0

US-10-029-345a-109 (1-665) x AR255944 (1-2377)

Qy 242 VALHSCYLeuAlaGlyIleSerArgSerAlaTrilleAlaIleAlaTyrIleMetLys 261
|||||
Db 864 GTCACATGCTGCTGGCATCTCCGCTCTGCACCATCGCATCGCCATCATGAG 923

RESULT 48 HSU27193 2377 bp mRNA linear PRI 08-DEC-1995
LOCUS HSU27193
DEFINITION Human protein-tyrosine phosphatase mRNA, complete cds.
ACCESSION U27193
VERSION U27193.1 GI:1109781
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2377)
AUTHORS Martelli, K.J., Seasholtz, A.F., Kwak, S.P., Clemens, K.K. and
Dixon, J.E.
TITLE hvH-5: a protein tyrosine phosphatase abundant in brain that
inactivates mitogen-activated protein kinase
JOURNAL J Neurochem. 65 (4), 1823-1833 (1995)

REFERENCE 2 (bases 1 to 2377)
AUTHORS Martelli, K.J.
TITLE Direct Submision
JOURNAL Submitted (16-MAY-1995) Karen J. Martelli, Biochemistry, University
of Michigan, M5416 Medical Science I, 1301 Catherine Street, Ann
Arbor, MI 48109-0606, USA
FEATURES Location/Qualifiers
source 1..2377
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hvH-5"
/issue_type="brain"
/dev_stage="fetal"
135..2012
/codon_start=1

FEATURES

source

1..2377
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hvH-5"
/issue_type="brain"
/dev_stage="fetal"
135..2012
/codon_start=1

CDS

/product="protein-tyrosine phosphatase"
/protein_id="AAA83151.1"
/db_xref="GI:1109782"
/translation="MAGDRLPRKMDAKKLASLRGGPGPIVDSRSEVYNSMHL
SSVNCSSKLVKRLQOKVTTIELIOPARSOVEETPDVVVDQSTRDASVLAAD
SELSILSLKIDGCPDSVALITGGFATPSSCFPGCLCGKPAALLPMSLSQCLPVSVG
LRIILPHLYIGSKVDLNDLMTONGISVYLNASNSCPDFICSRFWRVFINNYC
EKLLPMLDKSIEFIDKAKLSCOVIVHCLAGISRSATITAIAYIMKTMGSSDPAVEV
KDRRPSISPNFNLGOLLVEERTIKLLALLODPCPTSGTPPPSPAPAGAPLPLPP
PSESATGTGMAARBGSLSAGSGPPAPPTPPATSAOQGLRGLHLSDDLQDTNLRK
STSLDTKSAVPSRRPDGPPDPGAPFLCUDSPSGALGLSSPSPPAPAEAR
PPRRRRPPPPAGSPAPSPASISGLNGLDAAKQTPRRGLSALAPGLPGGQAPAGAM
APPLDSPGTSPSPDGWCFSPBGAQAGVLPAPFGAAGAPGPGGSDLRREAAAEAP
RDARTGMPEPAPETQFKRSCOMEPEGEWGEARAGEELALGKQASPSGSVEVTEV
S"

ORIGIN

Alignment Scores:
Pred. No.: 3.84e-09 Length: 2377
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.01% Indels: 0
DB: 9 Gaps: 0

US-10-029-345A-109 (1-665) x HSU27193 (1-2377)

Qy 242 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 261
DB 864 GTTCACTGTCTGGCTGGCATCTCTCGCTGCGCACCATCGCCATCGCTACATCATGAG 923

RESULT 49

LOCUS A59887 2415 bp DNA linear PAT 06-MAR-1998
DEFINITION Sequence 5 from Patent WO9706245.
ACCESSION A59887
VERSION A59887.1 GI:3715074
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Davies, K.E. and Theodosiou, A.
TITLE MAP KINASE PHOSPHATASE GENE AND USES THEREOF
JOURNAL Patent: WO 9706245-A 5 20-FEB-1997;
MEDICAL RES COUNCIL (GB)
FEATURES
source Location/Qualifiers
1..2415
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:
Pred. No.: 3.89e-09 Length: 2415
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.01% Indels: 0
DB: 6 Gaps: 0

US-10-029-345A-109 (1-665) x A59887 (1-2415)

Qy 242 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 261
DB 825 GTTCACTGTCTGGCTGGCATCTCTCGCTGCGCACCATCGCCATCGCTACATCATGAAA 884

RESULT 50

LOCUS A59888 2453 bp DNA linear PAT 06-MAR-1998
DEFINITION Sequence 6 from Patent WO9706245.
ACCESSION A59888
VERSION A59888.1 GI:3715075

KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Davies, K.E. and Theodosiou, A.
TITLE MAP KINASE PHOSPHATASE GENE AND USES THEREOF
JOURNAL Patent: WO 9706245-A 6 20-FEB-1997;
MEDICAL RES COUNCIL (GB)
FEATURES
source Location/Qualifiers
1..2453
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:
Pred. No.: 3.94e-09 Length: 2453
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.01% Indels: 0
DB: 6 Gaps: 0

US-10-029-345A-109 (1-665) x A59888 (1-2453)

Qy 242 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 261
DB 827 GTTCACTGTCTGGCTGGCATCTCTCGCTGCGCACCATCGCCATCGCTACATCATGAAA 886

Search completed: June 22, 2004, 09:06:14
Job time : 7621 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2004, 06:28:07 ; Search time 4394 Seconds

(without alignments)
4519.424 Million cell updates/sec

SUMMARIES

29: gb_gse2:*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Title: US-10-029-345A-109
Perfect score: 665
Sequence: 1 MAHEMIGTQIVTERLVALLE.....LGKVGSGSSFSGSGMEIIEVS 665
Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 27513289 seqs, 14931090276 residues
Word size: 1
Total number of hits satisfying chosen parameters: 200554
Minimum DB seq length: 1419
Maximum DB seq length: 200000000
Post-processing: Listing first 75 summaries

Command line parameters:
-MODE=frame+ p2n model -DEV=xlh
-Q/cgnt_1/USPTO.spool/US10029345/runat_21062004_164401_29765/app_query.fasta_1.839
-DB=EST -QPMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdt -LIST=75
-DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=1419 -MAXLEN=200000000
-USER=US10029345_@CGN_1_1_2810_@runat_21062004_164401_29765 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST: *
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gse_hum:*
18: em_gse_hiv:*
19: em_gse_pln:*
20: em_gse_vtc:*
21: em_gse_fun:*
22: em_gse_man:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_phg:*
27: em_gse_vtl:*
28: gb_gse61:*

Result No.	Score	Query Match	Length	DB ID	Description
1	67	10.1	3335	11	AK035652
2	42	6.3	2573	11	AK014441
3	20	3.0	2207	11	BC038231
4	12	1.8	2237	11	AK088665
5	12	1.8	2719	11	AK009131
6	12	1.8	2793	11	AK005062
7	12	1.8	2794	11	AK088468
8	9	1.4	1614	11	AK088059
9	9	1.4	1699	11	AK053746
10	9	1.4	2287	11	AK044726
11	9	1.4	2428	11	AK080964
12	8	1.2	1439	28	CC241307
13	8	1.2	1456	13	BU199156
14	8	1.2	1459	29	AG146840
15	8	1.2	1465	11	CNSLTI180
16	8	1.2	1502	12	BG121676
17	8	1.2	1603	28	BZ578841
18	8	1.2	1626	29	AY401953
19	8	1.2	1665	10	BF680537
20	8	1.2	1684	11	AY110929
21	8	1.2	1697	11	AK081315
22	8	1.2	1709	10	BE729007
23	8	1.2	1730	28	BZ576364
24	8	1.2	1742	12	BG859060
25	8	1.2	1746	29	AY400863
26	8	1.2	1770	12	BM320908
27	8	1.2	1792	12	BM321015
28	8	1.2	1839	11	AK045836
29	8	1.2	1867	11	AK017699
30	8	1.2	1918	11	AK075873
31	8	1.2	1919	29	CG755756
32	8	1.2	1965	11	AK047699
33	8	1.2	2008	28	BZ564814
34	8	1.2	2174	10	BF572336
35	8	1.2	2208	10	AW730576
36	8	1.2	2302	11	AK011786
37	8	1.2	2472	11	AY109354
38	8	1.2	2635	11	AK084572
39	8	1.2	2739	11	AK031373
40	8	1.2	3012	11	AK078652
41	8	1.2	3259	11	AK030955
42	8	1.2	3335	11	AK031191
43	8	1.2	3610	11	AK029006
44	8	1.2	3617	11	AK046988
45	8	1.2	3626	11	AK032373
46	8	1.2	3701	11	AK083783
47	8	1.2	3722	11	AK084916
48	8	1.2	3722	11	BC049163
49	8	1.2	3727	11	AK035810
50	8	1.2	3769	11	AK078517
51	8	1.2	4382	11	AK082310
52	8	1.2	4864	11	BC046246
53	8	1.2	1419	11	AK009999
54	7	1.1	1422	12	BG751601
55	7	1.1	1423	28	CC192711
56	7	1.1	1424	29	AG033031
57	7	1.1	1424	29	AY406613
58	7	1.1	1425	11	AK005246
59	7	1.1	1425	13	BQ214023
60	7	1.1	1425	29	AY415850
61	7	1.1	1426	28	CC189843
62	7	1.1	1426	28	CC231606

C 63	7	1.1	1427	13	B0058685	B0058685	AGENCOURT
C 64	7	1.1	1428	28	BZ553374	BZ553374	pacel-60
C 65	7	1.1	1429	28	CC241312	CC241312	CH261-36K
C 66	7	1.1	1429	28	CC246343	CC246343	CH261-360
C 67	7	1.1	1430	28	CC187178	CC187178	CH261-17G
C 68	7	1.1	1433	11	AK053824	AK053824	Mus muscu
C 69	7	1.1	1436	12	BM560037	BM560037	AGENCOURT
C 70	7	1.1	1437	29	AY406616	AY406616	Mus muscu
C 71	7	1.1	1438	11	AK040516	AK040516	Mus muscu
C 72	7	1.1	1440	29	CG754464	CG754464	F049-4-F0
C 73	7	1.1	1442	11	BC038834	BC038834	Homo sapi
C 74	7	1.1	1442	11	BC048999	BC048999	Homo sapi
C 75	7	1.1	1443	10	BF686707	BF686707	602143848

ALIGNMENTS

RESULT 1	AK035652	3325 bp	mRNA	linear	HTC 19-SRP-2003
LOCUS	AK035652				
DEFINITION	Mus musculus adult male urinary bladder cDNA, RIKEN full-length enriched library, clone:9530081F05 product:TRUNCATED MAPK. PHOSPHATASE 7 homolog [Homo sapiens], full insert sequence.				
ACCESSION	AK035652	GI:26330815			
VERSION	AK035652.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
JOURNAL	99279253				
MEDLINE	10349636				
PUBMED	11042159				
REFERENCE	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)				
JOURNAL	20499374				
MEDLINE	11042159				
PUBMED	11076861				
REFERENCE	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishimura, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	20530913				
MEDLINE	11076861				
PUBMED	11076861				
REFERENCE	4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)				
JOURNAL	5				
REFERENCE	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)				
JOURNAL	6				
REFERENCE	(bases 1 to 3325)				
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hitozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,				

COMMENT

Karich, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyu, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nomura, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toyu, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.c.riken.go.jp, URL:ftp://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.

Please visit our web site for further details.

URL:ftp://genome.gsc.riken.go.jp/

URL:http://fantom.gsc.riken.go.jp/

Location/Qualifiers

FEATURES

source

1..3325

organism="Mus musculus"

mol_type="mRNA"

strain="C57BL/6J"

db_xref="PANTOM DB:9530081F05"

db_xref="MG1:2359973"

db_xref="taxon:10090"

clone="9530081F05"

sex="male"

tissue_type="urinary bladder"

clone_id="RIKEN full-length enriched mouse cDNA library"

dev_stage="adult"

460..1080

note="unnamed protein product, TRUNCATED MAPK PHOSPHATASE 7 homolog [Homo sapiens] (SPTK1/Q96Q52, evidence: Fasta, 96.7%id, 86%length, match=369)

putative"

/codon_start=1

/protein_id="BAC29138.1"

/db_xref="GI:26330816"

/translation="MAHEMTGTQIVTSVALVLSGTEKVLIDSRPEVYNTSHILE AININCSLMRRLQODKVLITTELIOHSAKHKVIDCNQVVDQSDQVGSISDC FLVTILGLDERSFNSVHLIAGGFAPSRCPFLGCEKSTLPTCISQCLPVANIGPT RILNVLVIGCRDVLNVKASGKHLISLCSVASGAAVHYLSMIL"

CDS

ORIGIN

Alignment Scores:

Pred. No.: 1.86e-52 Length: 3325

Score: 67.00 Matches: 160

Percent Similarity: 96.39% Conservative: 0

Best Local Similarity: 96.39% Mismatches: 3

Query Match: 10.08% Indels: 6

DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x AK035652 (1-3325)

QY	15	LeuValAlaLeuLeuGluSerGlyThrcGluYuvValleuLeuileaspSerArgProPhe	34
DB	502	TTGGTGCTCTGCTGGAAAGTGAACGAAAGGCTGCTGAATTAATGAGCGACCAATT	561
QY	35	ValGluYrAenTrSerHisIleLeuGluAlaIleAenIleAenCySerLeuMet	54
DB	562	GTTGAATACATACGTCTCACATTTTGGAGCCATTAAATCAACTGCTCCAACTGATG	621
QY	55	LysArgArgLeuGlnGlnAspLysValleuileThrcGluLeuileGlnHisSerAlaLys	74
DB	622	AAGCGAAGTTGCAACAGGAAAGTATTAATTACAGAACTTATTCACCAATTCGCAAG	681

QY 75 HsLysValAspIleAspCysSer-GlnLysValValValTyrAspGlnSerSerGlnAs 94
682 CATTAAGGTGACATTCGTCGCA-TCAGAAAGGTGATGATTATGATCAAAAGTCCCAAG 740
DB 94 PValAla-SerLeuSerSerAspCysPheLeuThrValLeuLeuGlnLysLysLys- 113
741 TGTGTG-TTCTCTGTCGTGACAGCTGCTTCTCTCACTGACTCTCGGGAACAGAG-G 798
QY 114 SerPheAsnSerValHisLeuLeuAlaGlnGlyPheAlaGlnPheSerArgCysPhePro 133
799 AGTTTCAACTCTCTCCACTGCTGACGGGTGGCTTGAGTCTGCTGCTGCTTCTCCCT 858
QY 134 GlysLeuGlyGlnLysSerThrLeuValProThrCysIleSerGlnProCysLeuPro 153
859 GGGCTCTGTGAAGAAAGTCACTAGTCCCTACTGCAATCTGACGCTTCTTACT 918
DB 154 ValAlaAsnIleGlyProThrArgIleLeuProAsnLeuTyrLeuGlnCysGlnArg 173
919 GTTGCAGAACCTTGCGCACTCGAATCTTCCCAATCTATCTGCTGCGCAGAGAT 978
QY 174 ValLeuAsnLys 177
DB 979 GTCTCAACAG 990

RESULT 2
AK014441
LOCUS
DEFINITION
MUS musculus 18 days pregnant adult female placenta and extra embryonic tissue cDNA, RIKEN full-length enriched library, clone:3830417M1.7 product:RUNCATED MAPK PHOSPHATASE 7 homolog [Homo sapiens], full insert sequence.
ACCESSION
AK014441
VERSION
AK014441.1 GI:12852294
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
AUTHORS
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
REFERENCE
AUTHORS
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Taahito, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 685-690 (2001)
PUBMED
11076861
REFERENCE
AUTHORS
5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE
JOURNAL
REFERENCE
AUTHORS
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2573)

TITLE
JOURNAL
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-tes@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9216) Fax: 81-45-503-9216
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

COMMENT
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGGAGAGAGATCCAGACGCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 50.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGAGATCCAGATTAATTAATTCCTCCCTCCCTCCCT 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified plasmid KS(+) after bulk excision from Lambda FLC 1. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.
Location/Qualifiers
1. 2573
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:3830417M1.7"
/db_xref="MGI:1903116"
/db_xref="taxon:10090"
/clone="3830417M1.7"
/sex="female"
/tissue_type="placenta and extra embryonic tissue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18 days pregnant adult"
1. 2573
/note="RUNCATED MAPK PHOSPHATASE 7 homolog [Homo sapiens] (SPR1Q96Q2, evidence: FASTA, 96.7%id, 86%length, match=369)"

FEATURES
source
1. 2573
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:3830417M1.7"
/db_xref="MGI:1903116"
/db_xref="taxon:10090"
/clone="3830417M1.7"
/sex="female"
/tissue_type="placenta and extra embryonic tissue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18 days pregnant adult"

misc_feature
1. 2573
/note="RUNCATED MAPK PHOSPHATASE 7 homolog [Homo sapiens] (SPR1Q96Q2, evidence: FASTA, 96.7%id, 86%length, match=369)"

ORIGIN
Alignment Scores:
pred. No.: 1 22e-28 Length: 2573
Score: 42.00 Matches: 42
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.32% Indels: 0
DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x AK014441 (1-2573)

QY 624 LysArgSerCysGlnMetGlnPheGlnSerIleMetSerGlnAsnArgSerArg 643
DB 41 AAACGAGAGAGCTGCCAAATTCGATTTGAGAGAGAGATTAATGTCGAGAAACAGTCCAG 100

QY 644 GUGUleuGlytyValGlySerGlnSerSerPheSerGlySerMetGlnIleIleGlu 663
 DB 101 GAGGAGCTGGGCAAGGTGGGCAAGCCAGTCAGCTTCTCCGGGAGCATGAGATCATCGAG 160
 QY 664 ValSer 665
 DB 161 GTCTCT 166

RESULT 3
 BC038231 2207 bp mRNA linear HTC 01-OCT-2002
 LOCUS Homo sapiens, similar to dual specificity phosphatase 8, clone
 DEFINITION IMAGE:5547764, mRNA.
 ACCESSION BC038231 GI:233398534
 VERSION BC038231.1 GI:233398534
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2207)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (30-SEP-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
 COMMENT Contact: MGC help desk
 Email: cgaabcs-remail.nih.gov
 Tissue Procurement: ATCC/DCMD/DMP
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc_mgc@hri.nih.gov/
 Akher, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lalic, P., Legaepi, R.,
 Madanu, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
 Series: IRK Plate: 79 Row: a Column: 12
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4758211
 This clone has the following problem: frame shifted.

FEATURES
 source
 1. 2207
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5547764"
 /cissue_type="Skin, melanotic melanoma."
 /clone_lib="NIH MGC_72"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN
 Alignment Scores:
 Pldm. No.: 1.19e-07 length: 2207
 Score: 20.00 Matches: 20
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.01% Indels: 0
 DB: 11 Gaps: 0

US-10-029-345a-109 (1-665) x BC038231 (1-2207)

QY 222 ValHisCysLeuAlaGlyTlIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 261
 DB 861 GTCCACGTGCTGGCTGCATCTCCGCTCTGCACCATGCGCATGCGCTCATCATGTAAG 920

RESULT 4
 AK088665 2237 bp mRNA linear HTC 20-SEP-2003
 LOCUS Mus musculus 2 days neonate thymus thymic cDNA, RIKEN
 DEFINITION full-length enriched library, clone:B430023B14 product:dual
 specificity phosphatase 6, full insert sequence.
 ACCESSION AK088665 GI:26353717
 VERSION AK088665.1 GI:26353717
 KEYWORDS HTC; Cap trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE 4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 665-690 (2001)
 MEDLINE 11076861

REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 MEDLINE 11076861

REFERENCE 6
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akihira, S.,
 Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-Apr-2002) Yoshinide Hayashizaki, The Institute of

ACCESSION	AK009131	1
VERSION	AK009131.1	GI:12843734
KEYWORDS	HTC; CAP trapper.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-research@gs.c.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216).

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

```

1..2237
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:E430023B14"
/db_xref="MGI:247562"
/db_xref="taxon:10930"
/clone="E430023B14"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
464..1609
/note="unnamed protein product; dual specificity
phosphatase 6 (MD) [MGI:1914853, GB|NM_026268, evidence:
VLASTIN, 99%, match=8793]"

```

```

/codon_start=1
/protein_id="BAC00489.1"
/db_xref="GI:26353718"
/translation="MIDTLPVPVPPASPMALCKTVSYNLNQLELGNRLILLNDCKRQEI
YESSHSATINVAIPGIMTLRLPRLQKGIIPVALPTRCEDNDRFTRRCGIDTVLYKENS
PLDNENNGGSGEVLILKLLKDKDCAFAYLEGGSKFAQELPLHCEINLIDGSCSSHS
PLPLVAGNGRISLSDSSDLSLELDLPNPSATIDQSPSLNPSQSPPEVILLPLVYIGC
AKDSINIDVLEERIKIILNVNINLRLPEMAGEKTKROIPIEDMSQNLSPPEEA
SFIIDAGKNGCGVLAHCLASIVTVPLVLAQKILSMNADNYAVI VKKKSINISPNR
SFMDQILDFERTGLVSPCCNRRVPTDPLVFETTPSNQNVYDLSST"

```

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCES
AUTHORS

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia, Sclurognathi, Muridae, Murinae, Mus 1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10348636

TITLE

Ittoh, M., Kano, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Genome Res. 10 (10), 1617-1630 (2000)
20493374
11042159

TITLE

Shibata, K., Toyota, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, N., Kondo, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishibe, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneida, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Kozaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y. (2000) Integrated sequence analysis (ISA) system-184-format sequencing pipeline with 384 multicapillary sequencer genome Res. 10 (11), 1757-1771 (2000)

MEDLINE
PUBMED
REFERENCE
AUTHORS

The RIKEN genome Exploration Research Group Phase II Team and the PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

AUTHORS

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation

**JOURNAL
REFERENCE
AUTHORS**

Adachi, T., Iizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Himeno, K.,
Hisakawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, H., Izawa, M.,
Kasubawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numasaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D.,
Shibata, K., Shibata, Y., Shingawa, A., Shiraishi, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

JOURNAL

JOURNAL
Submitted (10-JULY-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama
Kanagawa 230-0045, Japan (E-mail: genome-research@gs.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>,
Fax: 81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

Enzyglopediadirect.com project of genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGCAGAGGAGAGATCCAGAGCTCTTTTCTTTTCTTTTATN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGCAGAGATTCCTCGATTATTAATTAATTCCTCCCTCCCTCC 3']. cDNA was cleaved

with XhoI and SctI. Cloning sites, 5' end: XhoI; 3' end: SctI.
Host: SDR.

FEATURES
source
Location/Qualifiers

1..2719
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:2310004A03"
/db_xref="MGI:1899068"
/db_xref="taxon:10090"
/clone="2310004A03"
/sex="male"
/tissue_type="tongue"
/clone_1fb="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
390..1535
/note="unnamed protein product; dual specificity
phosphatase 6 (MGD|MGI:1914853, GB|NM_026268, evidence:
BLASTN, 9%, match=2793)
putative"

CDS

/codon_start=1
/protein_id="BAB26093.1"
/db_xref="GI:12843735"
/translation="MTDLRPVPPFASMAICTVSGIANOLEIGNERFLIMDCRPOEL
YESSHIESAINVAIPGIMLRLOKGNLPPRALFTCEDRDFRRCGTPTVVLVDENS
SPWNTTGESVYGLILKLIKDEGCAFLKEGSKFOAEFALHGETMIDGSCSSSP
PVLVLGGLRISDSSDIESLDLPDSDGSPISNSOPSPVPEITLPELYIGC
ADOSTLDTLEERGIKYLNTVNTPLPENNAGEPKYKQIPISDMISQTLSPFPEAI
SPIDARGRKCCVYHCLAGISRSVTVTAYIMQKINLSWMDAYDIVKKKSNISPNF
NFMGLDPERLTGSSPCDNRVPPOLYFTTSPNONVYQVDSLOST"
2700..2705
/note="putative"
2719
/note="putative"
/note="putative"

ORIGIN

Alignment Scores:
Pred. No.: 6.78 Length: 2719
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservatave: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x AK009131 (1-2719)

Qy 240 ValieuValH:SCysleuNlaGlylleSsrArgSer 251

Db 1254 GTCTGTGTCATTGCTGTGACAGTATACAGCCGCTCT 1289

RESULT 6
AK005062 2793 bp mRNA linear HTC 20-SEP-2003
LOCUS
DEFINITION Mus musculus adult male liver cDNA, RIKEN full-length enriched
library, clone:1300019103 product:dual specificity phosphatase 6,
full insert sequence.

ACCESSION AK005062
VERSION AK005062.1 GI:12836727
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
1 Carninci, P. and Hayashizaki, Y.
AUTHORS
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multipipillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Submitted (10-JUL-2000) Yoshihide Hayashizaki. The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAGAGCGCCGCACTCCAGTCTTTTCTTTTCTTAA 3'], cDNA was
prepared by using trenaiose chemo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence [5'
GAGAGAGAGAGATCCAAAGCTCAATTAATTAATTAACCCGCCCC 3']. cDNA was
cleaved with XhoI and SctI. Cloning sites, 5' end: SctI; 3' end:
XhoI. Host: SDR.

FEATURES

source
Location/Qualifiers
1..2793
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:1300019103"
/db_xref="MGI:1897200"
/db_xref="taxon:10090"
/clone="1300019103"
/sex="male"

```

/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
463..1608
/notes="unnamed protein product; dual specificity
phosphatase 6 (MGD|MGI:1914853, GB|NM_026268, evidence:
BLASTN, 99%, match=2793)
putative"
/codon_start=1
/protein_id="BAB23786.1"
/db_xref="GI:12836728"
/translation="MIDTLRPVPFASMAICKTVSWLNEQLBENRLLIMDCRPOEL
YSSHIESAINVAIPGIMLRLOKGNLPVALFTPCGDRFRRCGTDIVVYDENS
SPMNTGGSVSVGLILKKIKDGCRAFYEGSPKQAPAFACETNLDGSSSSP
PLPVILGIGRISDSSSDIESPLDRPNATSDSPSPSPVELLPVLYGIC
AKDSTNDVLEEFGIKITLVATPNLPLFENAEFVKQIPIDHMSQNSQFPPEAI
SFIDEARGKGVVHCLAGISRSVTVVAVLMQKLNLSNDYDIDYMKKSNISNPF
NFMGDLDFERTIGLSPCDNRVPTPOLYFTFSPNQVYQVDSLQST"
2774..2779
/notes="putative"
2793
/notes="putative"
2793
/notes="putative"
2793

ORIGIN
polya_signal
polya_site

Alignment Scores:
Pred. No.:      6 98      length:      2793
Score:          12.00     Matches:      12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    1.80%      Indels:      0
DB:              11        Gaps:        0

US-10-029-345A-109 (1-665) x AK005062 (1-2793)
QY      240 ValLeuValHisCysLeuAlaGlyIleSerArgSer 251
Db      1327 GTCTGTGTCATGCTTGGCAGGATCAGCGGCTCT 1362

RESULT 7
AK088468
LOCUS      2794 bp      mRNA      linear      HTC 20-SEP-2003
DEFINITION Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN
full-length enriched library, clone:E430018G11 product:dual
specificity phosphatase 6, full insert sequence.
ACCESSION AK088468
VERSION   AK088468.1 GI:26353483
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
3
JOURNAL   RIKEN integrated sequence analysis (RISA) system--384-format
MEDLINE   sequencing pipeline with 384 multicapillary sequencer
PUBMED    Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
REFERENCE
4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
JOURNAL   The FANTOM Consortium and the RIKEN Genome Exploration Research
AUTHORS   Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2794)
REFERENCE Adachi,U., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
AUTHORS   Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiroaka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoh,H., Kawai,U., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohnato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Kanbara,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gs.c.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT  cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Truett/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
Source
location/Qualifiers
1..2794
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="PANTOM,DB:E430018G11"
/db_xref="MGI:2427365"
/db_xref="taxon:10090"
/clone="E430018G11"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
463..1608
/notes="unnamed protein product; dual specificity
phosphatase 6 (MGD|MGI:1914853, GB|NM_026268, evidence:
BLASTN, 99%, match=2793)
putative"
/codon_start=1
/protein_id="BAC40372.1"
/db_xref="GI:26353484"
/translation="MIDTLRPVPFASMAICKTVSWLNEQLBENRLLIMDCRPOEL
YSSHIESAINVAIPGIMLRLOKGNLPVALFTPCGDRFRRCGTDIVVYDENS
SPMNTGGSVSVGLILKKIKDGCRAFLYEGSPKQAPAFALHCTNLDGSSSSP

```

PLVILGIGLRISSSSSDIESLDRDPNSATDGSPLNSOPSPFVETILPVLIGC
 AKDSTNLDVLEBFGIKYLIANTPNLPLFNAGBFRKQIPIGHSWONSOFPPEAI
 SFIDEARGKNGCVLHCLAGISRSVTVAYIMKAKNLNMDADYIKAKKSNISFVF
 NEMQGLDFFERTLCLSSPCDNKRVTPOLYFTTSPNONVYDLSQST"
 2775..2780
 /note=".2780
 polyA_signal
 /note="putative"
 polyA_site
 /note="putative"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.98 Length: 2794
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.80% Indels: 0
 DB: 11 Gaps: 0
 US-10-029-345A-109 (1-665) x AK088468 (1-2794)
 Oy 240 ValIeuValH3CysLeuAlaGlyIleSerArgSer 251
 Db 1327 GTCTGTGATGCTTGTGGCAGGATACGCCGCTCT 1362
 RESULT 8 1614 bp mRNA linear HTC 20-SEP-2003
 AK088059 Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN
 LOCUS full-length enriched library, clone:E430002P08 product:dual
 DEFINITION specificity phosphatase 2, full insert sequence.
 ACCESSION AK088059.1 GI:26104661
 VERSION AK088059.1
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE
 1 Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Komno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper selected cDNAs to
 JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
 MEDLINE 20499374
 PUBMED 11042159
 3 Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,
 Komno,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,
 Sumi,N., Ishii,Y., Nakamura,S., Hamada,M., Nishine,T., Harada,A.,
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
 Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,B., Watabiki,M.,
 Yoneoka,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format
 JOURNAL sequencing pipeline with 384 multicapillary sequencer
 MEDLINE 20530913
 PUBMED 11076861
 4 The RIKEN Genome Exploration Research Group Phase II Team and the
 JOURNAL FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 REFERENCE Nature 409, 685-690 (2001)
 5 The FANTOM Consortium and the RIKEN Genome Exploration Research
 JOURNAL Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 1614)
 JOURNAL Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
 REFERENCE Hanagaki,T., Hara,A., Hashizume,W.,
 AUTHORS Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
 Hayashida,K., Hayatsu,N., Hiramoto,K., Hirakata,T., Hirozane,T.,
 Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Karukawa,T.,
 Kato,H., Kawai,D., Kojima,Y., Kondo,S., Komno,H., Kouda,M.,
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
 Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
 Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
 Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T.,
 Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akhiba,S.,
 Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
 Muramatsu,M. and Hayashizaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suenho-cho, Tsunumi-ku, Yokohama,
 Kanagawa 230-0045, Japan [E-mail:genome-res@gscc.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216]
 COMMENT cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
 Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
 Trust/MRC building Addenbrookes Hospital Cambridge) whose
 assistance we gratefully acknowledge.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.go.jp/
 URL:http://fantom.gsc.riken.go.jp/
 location/Qualifiers
 1..1614
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NOD"
 /db_xref="FANTOM DB:E430002P08"
 /db_xref="MG1:2426956"
 /db_xref="taxon:10090"
 /clone="E430002P08"
 /cell_type="thymic cells"
 /tissue_type="thymus"
 /clone_id="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="2 days neonate"
 93..1032
 /note="dual specificity phosphatase 2 (DB|U11330,
 evidence:BLASTN, 98%, match=1511)
 putative"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.9e+03 Length: 1614
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.35% Indels: 0
 DB: 11 Gaps: 0
 US-10-029-345A-109 (1-665) x AK088059 (1-1614)
 Oy 246 AlaGlyIleSerArgSerAlaThrIle 254
 Db 862 GCTGCATCTCTCGCTCAGCCACCATC 888
 RESULT 9 1699 bp mRNA linear HTC 20-SEP-2003
 AK053746 Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
 LOCUS library, clone:E130306H24 product:DUAL SPECIFICITY PROTEIN
 DEFINITION

PHOSPHATASE 4 (EC 3.1.3.48) (EC 3.1.3.16) (MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE-2) (MAP KINASE PHOSPHATASE-2) (MKP-2) homolog [Rattus norvegicus], full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK053746.1 GI:26343694
AK053746.1 GI:26343694
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

Carninci, P. and Hayashizaki, Y.

TITLE

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99279253

REFERENCE

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE

20499374

REFERENCE

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishii, K., Katsunari, T., Tanihara, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL

Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE

20530913

REFERENCE

AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE

Functional annotation of a full-length mouse cDNA collection

JOURNAL

Nature 409, 685-690 (2001)

MEDLINE

11076661

REFERENCE

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE

Functional annotation of a full-length mouse cDNA collection

JOURNAL

Nature 409, 685-690 (2001)

REFERENCE

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL

Nature 420, 563-573 (2002)

MEDLINE

11076661

REFERENCE

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hamaguchi, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirao, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, S., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numasaki, K., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Kanbara, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE

Direct Submission

JOURNAL

Submitted (16-JUL-2001)

REFERENCE

Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

Prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source

1. 1699

organism

/organism="Mus musculus"

mol_type

/mol_type="mRNA"

strain

/strain="C57BL/6J"

db_xref

/db_xref="FANTOM_DB:E130306H24"

db_xref

/db_xref="MGI:2425595"

taxon

/taxon="10090"

clone

/clone="E130306H24"

issue

/issue="eyeball1"

clone

/clone="E130306H24"

dev_stage

/dev_stage="0 day neonate"

note

/note="unlabeled protein product; DUAL SPECIFICITY PROTEIN PHOSPHATASE 4 (EC 3.1.3.48) (EC 3.1.3.16) (MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE-2) (MAP KINASE PHOSPHATASE-2) (MKP-2) homolog [Rattus norvegicus] (SWISSPROT|062767, evidence: FASTA, 98.5%id, 100%length, match=1194)

putative

/putative=""

protein

/protein_id="BAC35504.1"

db_xref

/db_xref="GI:26343695"

translation

/translation="MTWELTLEMDCSVLKRLNDRDENGSGAGSGSHGALGLSGCKVILDCRPLFSLHAGVIRGSVAVRNCNIVRRRAKSVSLQILPAEBEVRARL SGVSYAVYDRSPRAESLRSDSTVSLVQALRRARERDIDCLGGERFSEYPR FSKTKALAIIPVPVPTNESIDGSCSGPVLDDGGPVELPLVGLSYAARV DRMLDGLTALINVSDDCPNHPGQYVCIVEDNHRKADISGMPEALIEYIDVDC RGRVLVHCAGISRSRTTCLATLIMKRVRLBAEFVKQRSIISPNFSCQLQF EQQVLTSSAABASPSGPLRRERKATPTPTSPFVPSVGVHAPNSLPYLHSPIT TSPSC"

ORIGIN

Alignment Scores:

Pred. No.: 3.06e+03

Score: 9.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.35%

DB: 11

Gaps: 0

US-10-029-345A-109 (1-665) x AK053746 (1-1699)

OY 246

DB 972

RESULT 10

AK044726

LOCUS

DEFINITION

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

AK044726 2297 bp mRNA linear HTC 20-SEP-2003

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.,

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5

AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 2297)

AUTHORS Fuchida,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Adachi,J., Nishi,K., Hayashizaki,T., Hara,H., Hashizume,M., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozawa,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Katsukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES Retina RNA was provided by Dr. Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Please visit our web site for further details.

source URL: http://genome.gsc.riken.go.jp/ location/Qualifiers 1..2297

organism="Mus musculus"

mol_type="mrna"

strain="C57BL/6J"

db_xref="PANTOM,DB:A930037D12"

db_xref="MGI:2410058"

db_xref="taxon:10090"

clone="A930037D12"

tissue="retina"

clone_lib="RIKEN full-length enriched mouse cDNA library"

dev_stage="adult"

CDS

<1..571

/note="unnamed protein product; hypothetical protein (evidence: ProCrest,NCBI CDS Predictor) putative"

/codon_start=2

/protein_id="BAC32052.1"

/db_xref="GI:26336739"

/translation="GDSRATLTGRRAAARGRYKLTOLRQDPSRAGVLPPEPSG SPGRAGCGMRVGLGRGAECCGPTTGGVREGLHAYSGAPHPIDPVTPEP LGTSLERHAADLDLPRRGSQDPRRAAGTIGLRSGILGAPCSQAPGPPRAVQA HDGCTGGGPRDLGPGRAASSLSAUSL"

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Mismatches:	Indels:	Gaps:
Score: 4.22e+03	2297	9	0	0	0
Percent Similarity: 100.00%					
Best Local Similarity: 100.00%					
Query Match: 1.35%					
DB: 11					

US-10-029-345A-109 (1-665) x AK044726 (1-2297)

Qy 495 AlaGlnArgSerLeuLeuSerProleu 503

Db 1624 GCACAGAGAGAGCGCTTTGAGTCCTTTA 1650

RESULT 11

LOCUS AK080964

DEFINITION Mus musculus 4 days neonate male adipose cDNA, RIKEN full-length enriched library, clone: B430220P13 product: DUAL SPECIFICITY PROTEIN PHOSPHATASE 4 (EC 3.1.3.48) (EC 3.1.3.16) (MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE-2) (MAP KINASE PHOSPHATASE-2) (MKP-2) homolog (Rattus norvegicus), full insert sequence.

ACCESSION AK080964

VERSION AK080964.1 GI:26348914

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.,

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
Nature 409, 685-690 (2001)

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2428)

Naachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carrinchi,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numata,R., Ono,M., Ohnato,N., Okazaki,Y., Saito,K., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE
JOURNAL
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES
source
1. 2428
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FANTOM DB:B430220P13"
/db_xref="MGI:2412162"
/db_xref="taxon:10090"
/clone="B430220P13"
/sex="male"
/tissue_type="adipose"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="4 days neonate"
120. 1316
/note="unnamed protein product; DUAL SPECIFICITY PROTEIN PHOSPHATASE 4 (EC 3.1.3.48) (EC 3.1.3.16)
(MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE-2) (MAP KINASE PHOSPHATASE-2) (MKP-2) homolog [Rattus norvegicus] (SWISSPROT|Q62767, evidence: FASTA, 98.5%ID, 100%length, match=1194)
match=1194)
putative"
/codon_start=1
/protein_id="BAC38097.1"
/db_xref="GI:26348915"
/translation="MYTBLEREMDGSVLRKLNRPDNGGSGAGNGSGSHALGILSGGKCLLDRCRPLASAGITRGSVNRCTTYRRRAKGVSLBQIIPAEERARLRSGLYSAIVYDERSPRASLRSDSTVSLVQALRRNRERDILCKGYSRFSSEYEFCSKTKALAAIPVPVPTSTESLDLGGSCGTPHDGGFVEIIPLYLGSAYHAARDMLDALGITALLNVSDDCPNHFEGHYKCIPEVDNFKADISSWFEALYIDAVDCRGRLVHVGCGAIGRSATICTLAYLMKKRVLBEAFEVKORRSIIISNPSFMGLOLFESQVLTTSACAALASPSGRLRERKATPPTSGVFSPVSVGHAAPSLPLHSPI TTSPSC"
2404. 2409
/note="putative"
2428
/note="putative"

ORIGIN
polya_signal
polya_site
ORIGIN

Alignment Scores:
Pred. No.: 4,48e+03 Length: 2428
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 11 Gaps: 0

US-10-029-345a-109 (1-665) x AK080964 (1-2428)

Qy 246 AAGLYIISerzArgerAlarhile 254
Db 975 GCCGCATCTCCGATCAGCCACCATC 1001

RESULT 12
LOCUS CC241307/c 1439 bp DNA linear GSS 12-MAY-2003
DEFINITION CH261-10J7 Sp6.1 CH261 Gallus gallus genomic clone CH261-10J7,
genomic survey sequence.
ACCESSION CC241307
VERSION CC241307.1 GI:30567970
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1439)
AUTHORS Kremetzki,C., Higgindobham,J., Wylie,K., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTCACTATAG
Class: BAC ends
High quality sequence start: 205
High quality sequence stop: 561.
Location/Qualifiers

FEATURES
source
1. 1439
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-10J7"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - For library and clone ordering information: http://www.chori.org/bacpac"

ORIGIN
Alignment Scores:
Pred. No.: 2,32e+04 Length: 1439
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 28 Gaps: 0

US-10-029-345a-109 (1-665) x CC241307 (1-1439)

Qy 359 SerValProSerValProSerVal 366
Db 243 TCAGTTCCTTCTGTGCTTCAGTT 220

RESULT 13
BU199156

LOCUS BUI19156 1456 bp mRNA linear EST 05-SBP-2002
 DEFINITION DBCID04 DCB Homo sapiens cDNA, mRNA sequence.
 ACCESSION BUI19156
 VERSION BUI19156.1 GI:22717746
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1456)
 Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J., Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X., Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Cheng,Z., and Han,Z.
 Homo sapiens cDNA DCB clones
 Unpublished (2000)
 CONTACT: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn.
 FEATURES
 source
 1..1456
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_type="dendritic cells"
 /dev_stage="mature"
 /lab_host="BM25.8"
 /clone_lib="DCB"
 /note="Vector: pTritpLx2; Site_1: sfilA; Site_2: sfilB"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.35e+04 Length: 1456
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 13 Gaps: 0
 US-10-029-345A-109 (1-665) x BUI19156 (1-1456)
 QY 363 ValProSerValGlnProSerLeu 370
 Db 355 GTCCATCTGTCACCTCTCTTG 378
 RESULT 14
 AG146840 1459 bp DNA linear GSS 08-JUN-2002
 LOCUS AG146840
 DEFINITION Pan troglodytes DNA, clone: RP43-008124.TU, genomic survey sequence.
 ACCESSION AG146840
 VERSION AG146840.1 GI:16676518
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
 1
 Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Toeki,Y., Watanabe,H. and Sakaki,Y.
 BAC end sequences of Library RP43-008124
 Unpublished
 2 (bases 1 to 1459)
 Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Toeki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suenhiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library RP43-008124. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
 PRIMERS
 Sequencing: TU
 LIBRARY
 Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI
 Location/Qualifiers
 1..1459
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="RP43-008124.TU"
 /sex="male"
 /cell_type="lymphocytes"
 /clone_lib="RP43-008124 Chimpanzee Male BAC Library"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.36e+04 Length: 1459
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 29 Gaps: 0
 US-10-029-345A-109 (1-665) x AG146840 (1-1459)
 QY 486 SerValArgThrSerSerSerGly 493
 Db 387 AGGTCGACGTCGTCGAGTGGG 410
 RESULT 15
 CNSLTT18J 1465 bp mRNA linear HTC 18-JUN-2003
 LOCUS CNSLTT18J
 DEFINITION human full-length cDNA clone CS0DK001YK16 of HeLa cells of Homo sapiens (human).
 ACCESSION BX248296
 VERSION BX248296.1 GI:28207940
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1465)
 Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue
 2 (bases 1 to 1465)
 Genoscope.
 Direct Submission
 Submitted (04-FEB-2003) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 1st strand cDNA was primed with a NotI-oligo(4T) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 Location/Qualifiers
 1..1465
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK001YK16"
 /tissue_type="HeLa cells"

CDS
/note="Cot 25-normalized-vector pCMVSPORT_6"
<1..460
/note="unnamed protein product"
/codon_start=2
/protein_id="CAD62624.1"
/db_xref="GI:28207941"
/translation="SRLAGSGFRAPFPFRPPAPKVDAGLEAQRNPPLRSVAPLSPV
PKDYSAPORQCPQSNLSLKVMSEKSERVPPALSKSEAVSPILQRIQSRNSIMEA
AGGFQVHFHNSNDGRFRIERSADLEVISYSKSVTERKFKCTVLF"

ORIGIN

Alignment Scores:
Pred. No.: 2.37e+04 Length: 1465
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x CNSL718d (1-1465)

QY 415 AlaserleuHiglyPheSerSer 422
DB 723 GCCTCTCTACATGCTTTTCACG 700

RESULT 16
BG121676/c 1502 bp mRNA linear EST 30-JAN-2001
LOCUS 602352445F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:445067 5',
DEFINITION mRNA sequence.
ACCESSION BG121676
VERSION BG121676.1 GI:12615185
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1502)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM10236 row: j column: 12
High quality sequence stop: 65.
Location/Qualifiers
1..1502

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES

source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:445067"
/issue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 2.43e+04 Length: 1502
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.20% Indels: 0
DB: 12 Gaps: 0

US-10-029-345A-109 (1-665) x BG121676 (1-1502)

QY 383 LeuH16LeuSer1AaPaPgleu 390
DB 1343 CTCGATCTGTCGCGTCATCGTCTT 1320

RESULT 17

BZ578841/c 1603 bp DNA linear GSS 17-DEC-2002
LOCUS msh2.603.x1 msh Pseudomonas aeruginosa genomic clone msh2_603,
DEFINITION genomic survey sequence.
ACCESSION BZ578841
VERSION BZ578841.1 GI:27213902
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 1603)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: ckraymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1..1603

FEATURES

source
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/ecoln="WSH"
/db_xref="taxon:287"
/clone="msh2_603"
/clone_lib="msh"
/note="Environmental isolate. Whole genomic shotgun
library."

ORIGIN

Alignment Scores:
Pred. No.: 2.61e+04 Length: 1603
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 28 Gaps: 0

US-10-029-345A-109 (1-665) x BZ578841 (1-1603)

QY 7 GlyThGlnIleValThrGluArg 14
DB 1387 GGAACCTCAATTTGTACAGCGCA 1364

RESULT 18

AY401953/c 1626 bp DNA linear GSS 12-DEC-2003
LOCUS Pan troglodytes HCM1062 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY401953
VERSION AY401953.1 GI:39757939
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

Alignment Scores:
Pred. No.: 2.43e+04 Length: 1502
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

REFERENCE      1 (bases 1 to 1626)
AUTHORS        Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,
               Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
               Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smutsky,J.J.,
               Adams,M.D. and Cargill,M.
TITLE          Inferring nonneutral evolution from human-chimp-mouse orthologous
               gene trios
JOURNAL        Science 302 (5652), 1960-1963 (2003)
PUBMED         14671302
REFERENCE      2 (bases 1 to 1626)
AUTHORS        Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,
               Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
               Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smutsky,J.J.,
               Adams,M.D. and Cargill,M.
TITLE          Direct Submission
JOURNAL        Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
               Rockville, MD 20850, USA
COMMENT        This sequence was made by sequencing genomic exons and ordering
               them based on alignment.
FEATURES       location/Qualifiers
               source
               1..1626
               /organism="Pan troglodytes"
               /mol_type="genomic DNA"
               /db_xref="taxon:9598"
               <1..>1626
               /locus_tag="HCM1062"
ORIGIN
Alignment Scores:
Pred. No.:      2.65e+04      Length:      1626
Score:          8.00          Matches:      8
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    1.20%         Indels:         0
DB:             29            Gaps:           0
US-10-029-345A-109 (1-665) x AY401953 (1-1626)
QY             646 LeuGLYValGlySerGlnSer 653
Db             326 CTGGGTAAAGTGGGATCAACATCC 303
RESULT 19
LOCUS          BF680537               1665 bp      mRNA      linear      EST 21-DEC-2000
DEFINITION    6021552722F1 NIH_MGC_83 Homo sapiens CDNA clone IMAGE:4296425 5',
               mRNA sequence.
ACCESSION     BF680537
VERSION       BF680537.1 GI:11954432
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE     1 (bases 1 to 1665)
AUTHORS       NIH-MGC http://mgc.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgabbs-r@mail.nih.gov
               Tissue Procurement: CLONETECH Laboratories, Inc.
               cDNA Library Preparation: CLONETECH Laboratories, Inc.
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L1CM1148 row: 0 column: 18
               High quality sequence stop: 440.
               location/Qualifiers
               1..1665
               /organism="Homo sapiens"
               /mol_type="mRNA"
FEATURES       source

```

```

ORIGIN
Alignment Scores:
Pred. No.:      2.72e+04      Length:      1665
Score:          8.00          Matches:      8
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    1.20%         Indels:         0
DB:             10            Gaps:           0
US-10-029-345A-109 (1-665) x BF680537 (1-1665)
QY             498 SerLeuLeuSerProLeuHisArg 505
Db             1156 TCTGTGCTCTCTCCCTTCACCGT 1133
RESULT 20
LOCUS          AY110929               1684 bp      mRNA      linear      HTC 17-OCT-2002
DEFINITION    Zea mays CL2242_8 mRNA sequence.
ACCESSION     AY110929
VERSION       AY110929.1 GI:21215519
KEYWORDS      HTC.
SOURCE        Zea mays
ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE     1 (bases 1 to 1684)
AUTHORS       Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
               Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE         Maize Mapping Project/DuPont Consensus Sequences for Design of
               Overgo Probes
JOURNAL        Unpublished (2002)
REFERENCE     2 (bases 1 to 1684)
AUTHORS       Coe,E.H.
TITLE         Direct Submission
JOURNAL        Submitted (25-APR-2002) Maize Mapping Project, University of
               Missouri, Columbia, MO 65211, USA
COMMENT       If you are interested in getting corresponding physical clones,
               these are publicly available from ZmDB and may be found by BLAST
               searching at MSU, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
               www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
               maize cDNA sequences is either Virginia Walbot, Stanford or Pat
               Schnable, Iowa State, then clones may be requested from ZmDB:
               www.zmdb.iastate.edu.
               location/Qualifiers
               1..1684
               /organism="Zea mays"
               /mol_type="mRNA"
               /db_xref="MaizeDB:630884"
               /db_xref="taxon:4577"
               /clone_lib="Maize Mapping Project/DuPont Consensus
               Library"
               /note="this sequence is part of a project of EST
               assemblies resulting from the application of public
               contigs to seed DuPont contigs; this resource was
               assembled by DuPont as part of a collaboration for the

```

ORIGIN
overgo addressing of BACS in conjunction with the Maize Mapping Project"

Alignment Scores:
Pred. No.: 2.75e+04 Length: 1684
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x AY110929 (1-1684)

QY 549 SerThrProSerLeuThrSerSer 556
DB 427 TCGACTCGCTTGACGTCGCC 450

RESULT 21
AK081315 1697 bp mRNA linear HTC 20-SEP-2003
LOCUS AK081315/c
DEFINITION Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone: C130003P19 product: unclassifiable, full insert sequence.

ACCESSION AK081315
VERSION AK081315.1 GI:26099834
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Tameda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 111076861

JOURNAL MEDLINE 20530913
PUBMED 111076861
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arahawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashide, K., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hirozane, T.,

Horii, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnaka, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submision
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/

FEATURES
source
Location/Qualifiers
1..1697
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:C130003P19"
/db_xref="MGI:241366"
/db_xref="taxon:10090"
/clone="C130003P19"
/issue_type="head"
/issue_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days embryo"
1..1697
/note="unclassifiable"

misc_feature

ORIGIN
Alignment Scores:
Pred. No.: 2.77e+04 Length: 1697
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x AK081315 (1-1697)

QY 519 PheGlyLeuSerThrSerGlnGln 526
DB 1158 TTGGCTGCTGCAACCAACCA 1135

RESULT 22
BE729007 1709 bp mRNA linear EST 15-SEP-2000
LOCUS BE729007
DEFINITION 601562269F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3832024 5', mRNA sequence.
ACCESSION BE729007
VERSION BE729007.1 GI:10142999
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 1709)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: image.linl.gov
plate: LICMS08 row: 1 column: 17
High quality sequence stop: 30.

FEATURES

source

Location/Qualifiers
1..11709
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3832024"
/cissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_20"
/note="Organ: Skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(C). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:

Pred. No.:	2.79e+04	Length:	1709
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	10	Gaps:	0

US-10-029-345A-109 (1-665) x BE729007 (1-1709)

QY 644 GluGluLeuGlyValGlySer 651

Db 216 GAGGATTGGAAGTGGCTTC 239

RESULT 23

BZ576364 1730 bp DNA linear GSS 17-DEC-2002
BZ576364
LOCUS msh2.4896.v2 msh Pseudomonas aeruginosa genomic clone msh2_4896,
DEFINITION genomic survey sequence.
ACCESSION BZ576364
VERSION BZ576364.1 GI:27211425
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 1730)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,D.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Classes: shotgun.
Location/Qualifiers
1..1730
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"

FEATURES

source

1..1730
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"

/clone="msh2_4896"
/clone_lib="msh"
/note="Environmental isolate. Whole genomic shotgun library."

ORIGIN

Alignment Scores:

Pred. No.:	2.83e+04	Length:	1730
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.20%	Indels:	0
DB:	28	Gaps:	0

US-10-029-345A-109 (1-665) x BZ576364 (1-1730)

QY

580 SerAlaTYSerCySserGlnLeu 587

RESULT 24

BG859060 1742 bp mRNA linear EST 29-MAY-2001
LOCUS 1024061B08.x1 C. reinhardtii CC-1690, normalized, Lambda Zap II
DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BG859060
VERSION BG859060.1 GI:14240244
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonae.
1 (bases 1 to 1742)
Grossman,A., Davies,J., Federpiel,N., Harris,E., Lefebvre,P.,
McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants: project phase 2
Unpublished (2000)
CONTACT: Charles Hauser
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
Location/Qualifiers
1..1742
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO2 and HS medium bubbled with 5% CO2. POLYA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExSassist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

FEATURES

source

1..1742
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO2 and HS medium bubbled with 5% CO2. POLYA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExSassist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

ORIGIN

Alignment Scores:

Pred. No.:	2.85e+04	Length:	1742
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 12 Gaps: 0

US-10-029-345a-109 (1-665) x BG8559060 (1-1742)

Qy 473 Thralaarpseraepsergin 480
 Db 106 ACCGACGCGCCGACGACTCTCAG 129

RESULT 25
 AY400863 1746 bp DNA linear GSS 12-DEC-2003
 LOCUS Homo sapiens HCM0699 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY400863
 VERSION AY400863.1 GI:39756852
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1746)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
 Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Snihsy,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 TITLE This sequence was made by sequencing genomic exons and ordering
 COMMENT chem based on alignment.
 FEATURES
 source
 1. 1746
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..21746
 /locus_tag="HCM0699"

Alignment Scores:
 Pred. No.: 2.86e+04 Length: 1746
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 29 Gaps: 0

US-10-029-345a-109 (1-665) x AY400863 (1-1746)

Qy 498 SerleulenserProleuHsargy 505
 Db 687 AGCCTTCTTCACCACTTCATCGC 710

RESULT 26
 BM320908 1770 bp mRNA linear EST 03-JAN-2002
 LOCUS rockefeller.0.435 Mastigamoeba balamuthi lambda ZAP II library
 DEFINITION Mastigamoeba balamuthi cDNA similar to actin-binding protein
 (coronin), mRNA sequence.
 ACCESSION BM320908
 VERSION BM320908.1 GI:18055314
 KEYWORDS EST.

SOURCE Mastigamoeba balamuthi
 ORGANISM Mastigamoeba balamuthi
 Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
 REFERENCE 1 (bases 1 to 1770)
 Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Senses,C.W.,
 Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
 Philippe,H.
 TITLE The analysis of 100 genes supports the grouping of three highly
 JOURNAL divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
 MEDLINE Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
 PUBMED 21819461
 CONTACT: Muller Miklos
 COMMENT Laboratory of Biochemical Parasitology
 The Rockefeller University
 1230 York Avenue, New York, NY 10021, USA
 Email: mmuller@rockvax.rockefeller.edu
 Insert Length: 1770 Std Error: 0.00
 POLYA=Yes.

FEATURES
 source
 1. 1770
 /organism="Mastigamoeba balamuthi"
 /mol_type="mRNA"
 /strain="ATCC 30984"
 /db_xref="taxon:108607"
 /clone_id="Mastigamoeba balamuthi lambda ZAP II library"
 /note="syn: Phreatamoeba balamuthi"

Alignment Scores:
 Pred. No.: 2.9e+04 Length: 1770
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 12 Gaps: 0

US-10-029-345a-109 (1-665) x BM320908 (1-1770)

Qy 488 ArgThSerSerSerg1YthAla 495
 Db 1324 AGAACGACGAGCTCAGAACCGCG 1347

RESULT 27
 BM321015 1792 bp mRNA linear EST 03-JAN-2002
 LOCUS rockefeller.0.586 Mastigamoeba balamuthi lambda ZAP II library
 DEFINITION Mastigamoeba balamuthi cDNA similar to actin-binding protein
 (coronin), mRNA sequence.
 ACCESSION BM321015
 VERSION BM321015.1 GI:18055421
 KEYWORDS EST.
 SOURCE Mastigamoeba balamuthi
 ORGANISM Mastigamoeba balamuthi
 Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
 REFERENCE 1 (bases 1 to 1792)
 Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Senses,C.W.,
 Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
 Philippe,H.
 TITLE The analysis of 100 genes supports the grouping of three highly
 JOURNAL divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
 MEDLINE Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
 PUBMED 21819461
 CONTACT: Muller Miklos
 COMMENT Laboratory of Biochemical Parasitology
 The Rockefeller University
 1230 York Avenue, New York, NY 10021, USA
 Email: mmuller@rockvax.rockefeller.edu
 Insert Length: 1792 Std Error: 0.00
 POLYA=Yes.

FEATURES
 source
 1. 1792
 /organism="Mastigamoeba balamuthi"

```

/organism="Mastigamoeba balamuthi"
/mol_type="mRNA"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_id="Mastigamoeba balamuthi lambda ZAP II library"
/notes="syn: Phreatamoeba balamuthi"

ORIGIN

Alignment Scores:
Pred. No.:      2.94e+04      Length:      1792
Score:          8.00         Matches:      8
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     1.20%       Indels:      0
DB:              12         Gaps:        0

US-10-029-345A-109 (1-665) x BM321015 (1-1792)

QY      488 ArgThrsErgserGlyThrAla 495
Db      1346 AGAACGACGAGCTCAGCAACCGG 1369

RESULT 28
AKO45836
LOCUS   1839 bp      mRNA      linear      HTC 20-SEP-2003
DEFINITION Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B230313L14 product:unknown EST, full insert sequence.
ACCESSION AKO45836
VERSION   AK045836.1 GI:26091148
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS   Carninci,P. and Hayashizaki,Y.
TITLE     High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED    10349636

REFERENCE
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED    11042159

REFERENCE
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE     RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED    11076861

REFERENCE
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE     Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
AUTHORS   5
TITLE     The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL   Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
NATURE    Nature 420, 563-573 (2002)

```

```

REFERENCE
AUTHORS   6 (bases 1 to 1839)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1..1839
organism="Mus musculus"
mol_type="mRNA"
strain="C57BL/6J"
db_xref="FANTOM DB:B230313L14"
/db_xref="MG1:2411215"
/db_xref="taxon:10090"
/clone="B230313L14"
sex="male"
tissue_type="corpora quadrigemina"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..1839
note="unknown EST (GB|AA689954, evidence: BLASTN, 99%, match=358)"

misc_feature

ORIGIN

Alignment Scores:
Pred. No.:      3.02e+04      Length:      1839
Score:          8.00         Matches:      8
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     1.20%       Indels:      0
DB:              11         Gaps:        0

US-10-029-345A-109 (1-665) x AKO45836 (1-1839)

QY      171 GlnArgAspValIleuAbnLysGlu 178
Db      275 CAGCGAGATGTCTGAAACAAAGAA 298

RESULT 29
AKO17699
LOCUS   1867 bp      mRNA      linear      HTC 20-SEP-2003
DEFINITION Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:573047606 product:DU1069P2.2 (OUTER MITOCHONDRIAL MEMBRANE TRANSLOCASE HYPO34P) (SIMILAR TO TRANSLOCASE OF OUTER MITOCHONDRIAL MEMBRANE 34) (HYPOTHETICAL 34.6 KDa PROTEIN) (UNKNOWN) (PROTEIN FOR MGC:3515) homolog (Homo sapiens), full insert sequence.
ACCESSION AKO17699
VERSION   AKO17699.1 GI:12857072
KEYWORDS  HTC; CAP trapper.

```

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
PUBMED	99479253
REFERENCE	10349686
AUTHORS	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
PUBMED	20499374
REFERENCE	11042159
AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, Y., Nishi, K., Kitasuna, T., Teshiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, N., Nishino, T., Harada, A., Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubiki, M., Yoneda, Y., Ichikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE	RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
PUBMED	20530913
REFERENCE	11076861
AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
PUBMED	
REFERENCE	5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 1867)
JOURNAL	Adachi, J., Aizawa, K., Akabira, S., Akimura, T., Arai, A., Aono, H., Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Fukushima, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kaushika, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koys, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
PUBMED	
REFERENCE	Direct Submission
AUTHORS	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome and Chemical Sciences Center, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsutsui-Ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel. 81-45-503-9222, Fax: 81-45-503-9216]
TITLE	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details
JOURNAL	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGCAGACAAAGATCCAGACGCTTTTTTTTNN 3'], cDNA was prepared by using trihaloase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went
COMMENT	

through one round of subreaction to Rot = 100.0. Second strand cDNA was prepared with the primer adaptor of sequence [5' GAGAGGACATTCGACGTTAATTAAATATCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pluscript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamI. Host: DH10B.

Location/Qualifiers

1. 1867

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:5730476E06"

/db_xref="MGJ:1838741"

/db_xref="taxon:10090"

/clone="5730476E06"

/tissue_type="whole body"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="8 days embryo"

76. 1005

/note="unnamed protein product; DJ1069P2.2 (OUTER MITOCHONDRIAL MEMBRANE TRANSLOCASE HTOM34P) (SIMILAR TO TRANSLOCASE OF OUTER MITOCHONDRIAL MEMBRANE 34) (HYPOTHETICAL 34.6 KDA PROTEIN) (UNKNOWN) (PROTEIN FOR MGC:3515) homolog (Homo sapiens) (SPRR109NT23, evidence: PASTY, 87.1%ID, 100%length, match=927)

putative"

/codon_start=1

/protein_id="BAB30882.1"

/db_xref="GI:12657073"

/translation="MAPKVSDFVEQLRAAGNPNRNGVGEASALYERALLRLOAGSS ADPEESVLYSNRAACTLCKNGCTDCTSLALVPISIKELRLPRASVYELKETA LAVADVXTVLQIDNSVASALEGINRITRALMDSLGEPWRKLPPIVVPVSAGKRNLS LPDNNHETAKTKSKKATKSRVSPSGDVERAKLKEBENDLVKGNHKA1EKSE SLICSLSEATKTYENRALCHLVLRQYEAVDCTEALKLDSKNVKAIFYRRAQAVKALKD YKSLSDISLSLQIEPRNGAQLRQEVNQNMM"

1848. 1853

/note="putative"

1867

/note="putative"

1867

/note="putative"

Alignment Scores:

Pred. NO.:

Score: 3.07e+04 Length: 1867

8.00 Matches: 8

100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.20% Indels: 0

DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x AK017699 (1-1867)

Qy 384 HisLenSerAlaAspArgLeuGlu 391

|||||

Db 1421 CACCTGAGTGTGACAGCCTAGAA 1444

RESULT 30

AK075873

LOCUS

DEFINITION Mus musculus adult male tongue cDNA, Riken full-length enriched library, clone:2310009L17 product:DJ1069P2.2 (OUTER MITOCHONDRIAL MEMBRANE TRANSLOCASE HTOM34P) (SIMILAR TO TRANSLOCASE OF OUTER MITOCHONDRIAL MEMBRANE 34) (HYPOTHETICAL 34.6 KDA PROTEIN) (UNKNOWN) (PROTEIN FOR MGC:3515) homolog (Homo sapiens), full insert sequence.

AK075873

ACCESSION

VERSION AK075873.1 GI:26344741

KEYWORDS

SOURCE

ORGANISM Mus musculus (house mouse)

ECUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

AUTHORS

TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20493974
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kikunai, T., Tashiro, H., Itoh, M., Sumi, N., Iehi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
MEDLINE 11076861
PUBMED 11076861
REFERENCE 5
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 11076861
PUBMED 11076861
REFERENCE 6
AUTHORS Adachi, U., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 220-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
location/Qualifiers
1. 1918
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:2310009L17"
/db_xref="MGI:1895159"
/db_xref="taxon:10090"

CDS
/clone="2310009L17"
/sex="male"
/tissue_type="longue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
128..1057
/note="unnamed protein product; D1069P2.2 (OUTER MITOCHONDRIAL MEMBRANE TRANSLOCASE HTOM34P) (SIMILAR TO TRANSLOCASE OF OUTER MITOCHONDRIAL MEMBRANE 34) (HYPOTHETICAL 34.6 KDA PROTEIN) (UNKNOWN) (PROTEIN FOR MGC:3515) homolog [Homo sapiens] (SPTR|Q9NTZ3, evidence: FASTA, 87.1%ID, 100%length, match=927)
putative"
/codon_start=1
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLYENRACVYLKDCGCTDCKDTSALALVPSIKPLLRASAYEALERKALAYDYKTVLIDNSVASALGIRITRALMDSIGPERKLTPIIPVPSAQGRMSLPDNDKETAETKSKREATKSRVPSAGDERALALKEGNDLVKGNHKKALIKYSLSLSSLSATVSRNALCHLVLKQKKAVDCTEALKLDGNVAFYRRAQYVALDYKSSLSIDISLILQIEPRNGPAOKLQREVNQNMN"
1901..1906
/note="putative"
/protein_id="BAC36020.1"
/db_xref="GI:26344742"
/translation="MAPVSPSVBQLRAQNGNFRNGOYGEASALYERALTLLQARSDADPEESLY


```

/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="pPa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

```

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
3.16e+04	1919	8.00	0	0	0	0
Percent Similarity:	100.00%					
Best Local Similarity:	100.00%					
Query Match:	1.20%					
DB:	29					

US-10-029-345A-109 (1-665) x CG755756 (1-1919)

QY 548 ThSerrThProSerLeuThSer 555

Db 29 ACCAGCACCCCTCGCTGACATCA 6

RESULT 32

AK047699/c

mus musculus adult male corpus striatum cDNA, RIKEN full-length enriched library, clone: C030013H07 product: unclassifiable, full insert sequence.

ACCESSION

AK047699

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1965)

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohnaka, N., Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE

Direct Submision Submitted (16-JUN-2001) Yoshitake Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://phantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source

1..1965
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:C030013H07"
/db_xref="MGI:2413445"
/db_xref="taxon:10090"
/clone="C030013H07"
/sex="male"
/cisse_type="corpus striatum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..1965
/note="unclassifiable"

ORIGIN

misc_feature

Alignment Scores:	Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	3.24e+04	1965	8.00	0	0	0	0
Percent Similarity:	100.00%						
Best Local Similarity:	100.00%						
Query Match:	1.20%						
DB:	11						

US-10-029-345A-109 (1-665) x AK047699 (1-1965)

QY 564 SerHapherYSerHaseRala 571

Db 68 AGTCATTATTCTCTGCTAGTCA 45

RESULT 33

BZ564814/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

BZ564814 2008 bp DNA linear GSS 17-DEC-2002
pac62-164_4936.y2 pac62-164 Pseudomonas aeruginosa genomic clone
pac62-164_4936, genomic survey sequence.
BZ564814 BZ564814.1 GI:27191372
GSS.
Pseudomonas aeruginosa
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 2008)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
JOURNAL Pseudomonas aeruginosa library
COMMENT J. Bacteriol. (2002) in press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1. 2008
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pac2-164_4936"
/clone_1lb="pac2-164"
/note="clinical isolate 2-164 whole genomic shotgun
library."

Alignment Scores:
Pred. No.: 3.32e+04 Length: 2008
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 28 Gaps: 0

US-10-029-345A-109 (1-665) x BZ564814 (1-2008)

QY 357 ProAlaServAlProServAlPro 364
1372 CCGGCTCTGTCTCCCTCGGTCGN 1349

RESULT 34
BF572336/c 2174 bp mRNA linear EST 12-DEC-2000
LOCUS 60207782F1 NIH_MGC_62 Homo sapiens CDNA clone IMAGE:4252004 5',
DEFINITION mRNA sequence.
ACCESSION BF572336 2174 bp mRNA linear EST 12-DEC-2000
VERSION BF572336 Homo sapiens CDNA clone IMAGE:4252004 5',
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2174)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCDN/DTF
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/MLNI at:
http://image.llnl.gov
Plate: LCM1076 row: d column: 21
High quality sequence start: 4
High quality sequence stop: 26.
Location/Qualifiers
1. 2174
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source

/clone="IMAGE:4252004"
/tissue_type="melanotic melanoma, high MDR"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1lb="NIH MGC_62"
/note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctggcgc); Site_2: SfiI (ggccataggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGCGGCACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."

ORIGIN
Alignment Scores:
Pred. No.: 3.61e+04 Length: 2174
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 10 Gaps: 0

US-10-029-345A-109 (1-665) x BF572336 (1-2174)

QY 323 ValProAlaValSerGluGly 330
681 GTGCTGCTGTGTGAGGCGCGG 658

ORIGIN

US-10-029-345A-109 (1-665) x AK011786 (1-2302)

QY 384 HisLeuSer1AAspArgLeuGlu 391
 |||||
 Db 1855 CACCTGAGTGTGACAGGCTAGAA 1878

RESULT 37
 AY109354/c 2472 bp mRNA linear HTC 17-OCT-2002
 LOCUS
 DEFINITION Zea mays C1199_1 mRNA sequence.
 ACCESSION AY109354
 VERSION AY109354.1 GI:21213009
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 2472)
 Hailey,C.F., Dolan,M., Mao,G.H., Vogel,J.M., Whitesitt,M.S.,
 Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
 Overgo Probes
 JOURNAL Unpublished (2002)
 REFERENCE 2 (bases 1 to 2472)
 AUTHORS Coe,E.H.
 JOURNAL Direct Submission
 REFERENCE Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 COMMENT If you are interested in getting corresponding physical clones,
 these are publicly available from ZmDB and may be found by BLAST
 searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat
 Schnable, Iowa State, then clones may be requested from ZmDB:
 www.zmdb.iastate.edu.

FEATURES
 source Location/Qualifiers
 1..2472
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:4577"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Consensus
 library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed DuPont contigs; this resource was
 assembled by DuPont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

ORIGIN
 Alignment Scores:
 Pred. No.: 4.14e+04 Length: 2472
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x AY109354 (1-2472)

QY 486 SerValArgThrSerSerGly 493
 |||||
 Db 2090 TCTGTAGAACTTCATCTCTGGG 2067

RESULT 38
 AK084572/c 2635 bp mRNA linear HTC 20-SEP-2003
 LOCUS
 DEFINITION Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched
 library, clone: D310016F04 product: unknown EST, full insert
 sequence.
 ACCESSION AK084572

VERSION AK084572.1 GI:26102078
 KEYWORDS HTC; CAP trapper
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 2 99279253
 10349636

AUTHORS
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
 MEDLINE Genome Res. 10 (10), 1617-1630 (2000)
 PUBMED 20499374
 11042159

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
 Komno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
 Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,S., Kawai,J.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 2635)

JOURNAL Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
 Fukuda,S., Furuno,M., Hanegaki,T., Hara,A., Hashizume,W.,
 Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T.,
 Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
 Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Komno,H., Kouda,M.,
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
 Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
 Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
 Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T.,
 Sogabe,Y., Tagami,M., Tagawa,A., Takaishi,F., Takaku-Akahira,S.,
 Takeda,Y., Tanaka,T., Tomaru,A., Toyu,T., Yasunishi,A.,
 Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission
 JOURNAL Submitted (16-APR-2002) Yoshinide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.go.jp,
 URL:htp://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216]

COMMENT CDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL:htp://genome.gsc.riken.go.jp/
 URL:htp://fantom.gsc.riken.go.jp/.

```

FEATURES
  source
    1. .2635
      /location/Qualifiers
      /mol_type="Mus musculus"
      /strain="C57BL/6J"
      /db_xref="PANTOM.DB:D330016F04"
      /db_xref="MGI:2431237"
      /db_xref="taxon:10090"
      /clone="D330016F04"
      /tissue_type="heart"
      /clone_lib="RIKEN full-length enriched mouse cDNA library"
      /dev_stage="13 days embryo"
      1. .2635
      /note="unknown EST (GB|BB012554, evidence: BLASTN, 99%,
      match=495)"

ORIGIN

Alignment Scores:
Pred. No.:      4.43e+04      Length:      2635
Score:          8.00         Matches:      8
Percent Similarity: 100.00%   Conservative: 0
Beet Local Similarity: 100.00% Mismatches:      0
Query Match:     1.20%       Indels:          0
DB:              11          Gaps:           0

US-10-029-345A-109 (1-665) x AK084572 (1-2635)

QY      526 GlnHleuThrySeraIagly 533
      |||||
      204 CAACATCTGACCAAGTCAGCTGCG 181

RESULT 39
AK031373
LOCUS   2739 bp      mRNA      linear      HTC 18-SEP-2003
DEFINITION Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
            enriched library, clone:6030414M07 product:zinc finger protein 91,
            full insert sequence.
ACCESSION AK031373
VERSION   AK031373.1 GI:26082294
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS   Carninci,P. and Hayashizaki,Y.
TITLE     High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED    10349636

REFERENCE
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
          Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED    11042159

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
  
```

```

TITLE     FANTOM Consortium.
JOURNAL   Functional annotation of a full-length mouse cDNA collection
REFERENCE Nature 409, 685-690 (2001)
AUTHORS   5

TITLE     The PANTOM Consortium and the RIKEN Genome Exploration Research
JOURNAL   Group Phase I & II Team.
REFERENCE Analysis of the mouse transcriptome based on functional annotation
AUTHORS   of 60,770 full-length cDNAs
          Nature 420, 563-573 (2002)
          6 (bases 1 to 2739)
          Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
          Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
          Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoaka,T., Hirozane,T.,
          Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanakawa,T.,
          Katoh,H., Kawai,U., Kojima,Y., Kondo,S., Kono,H., Kouda,M.,
          Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
          Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohnato,N.,
          Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
          Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
          Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaki,Akashira,S.,
          Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
          Muramatsu,M. and Hayashizaki,Y.
          Direct Submission
          Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
          Physical and Chemical Research (RIKEN), Laboratory for Genome
          Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
          Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp,
          URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
          Fax:81-45-503-9216)

COMMENT   cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
          Please visit our web site for further details.
          URL:http://genome.gsc.riken.go.jp/
          URL:http://fantom.gsc.riken.go.jp/.

FEATURES
  source
    1. .2739
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6J"
      /db_xref="PANTOM.DB:6030414M07"
      /db_xref="MGI:2394592"
      /db_xref="taxon:10090"
      /clone="6030414M07"
      /sex="male"
      /tissue_type="testis"
      /clone_lib="RIKEN full-length enriched mouse cDNA library"
      /dev_stage="13 days embryo"
      1. .2739
      /note="zinc finger protein 91 (locuslink109910,
      GB|U05342, evidence: BLASTN, 99%, match=838) "

ORIGIN

misc_feature
  1. .2635
    /location/Qualifiers
    /mol_type="Mus musculus"
    /strain="C57BL/6J"
    /db_xref="PANTOM.DB:6030414M07"
    /db_xref="MGI:2394592"
    /db_xref="taxon:10090"
    /clone="6030414M07"
    /sex="male"
    /tissue_type="testis"
    /clone_lib="RIKEN full-length enriched mouse cDNA library"
    /dev_stage="13 days embryo"
    1. .2635
    /note="zinc finger protein 91 (locuslink109910,
    GB|U05342, evidence: BLASTN, 99%, match=838) "

Alignment Scores:
Pred. No.:      4.62e+04      Length:      2739
Score:          8.00         Matches:      8
Percent Similarity: 100.00%   Conservative: 0
Beet Local Similarity: 100.00% Mismatches:      0
Query Match:     1.20%       Indels:          0
DB:              11          Gaps:           0

US-10-029-345A-109 (1-665) x AK031373 (1-2739)

QY      414 AaaAlaSerLeuHieglYpSer 421
      |||||
      1801 GCAGCATCCCTCCATGCTTCTTC 1824

RESULT 40
AK078652/c
LOCUS    AK078652      3012 bp      mRNA      linear      HTC 18-SEP-2003
  
```

DEFINITION Mus musculus adult male adrenal gland cDNA, RIKEN full-length enriched library, clone:7330435N05 product:B-cell translocation gene 3, full insert sequence.

ACCESSION AK078652

VERSION AK078652.1 GI:26098017

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Mech. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

MEDLINE 11076861

PUBMED 11076861

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 11076861

PUBMED 11076861

REFERENCE 6

AUTHORS Adachi, U., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akihira, S., Takebe, Y., Tanaka, T., Toyama, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suenitro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to

Prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source

1..3012

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM,DB:7330435N05"

/db_xref="MGI:2396891"

/db_xref="taxon:10090"

/clone="7330435N05"

/sex="male"

/tissue_type="adrenal gland"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

1..3012

/note="B-cell translocation gene 3 (MGD|MGI:109532, GB|NM_009770, evidence: BLASTN, 98%, match=438)"

misc_feature

Alignment Scores:

Pred. No.: 5.11e+04

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.20%

DB: 11

Gaps: 0

US-10-029-345a-109 (1-665) x AK078652 (1-3012)

QY 575 GYSerAlaSerTYrSerAlaTYr 582

Db 2395 GGCTCCGATCATCTCGCTTAC 2372

RESULT 41

AK030955

LOCUS

DEFINITION AK030955 3259 bp mRNA linear HTC 18-SEP-2003

Mus musculus adult male thymus cDNA, RIKEN full-length enriched library, clone:5830469C09 product:similar to RAB3-GAP REGULATORY DOMAIN [Homo sapiens], full insert sequence.

ACCESSION AK030955

VERSION AK030955.1 GI:26082074

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Mech. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 3359)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanehaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, P., Imocani, K., Iehli, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunataki, R., Ohno, M., Ohsumi, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takanehhi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama Kanagawa 230-0045, Japan [E-mail:genome-res@gscc.riken.go.jp, URL: http://genome-gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216]
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome-gsc.riken.go.jp/ URL: http://lancom.gec.riken.go.jp/. Location/Qualifiers
FEATURES	source
	1..3259
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM_DB:5830469C09"
	/db_xref="MGI:2394118"
	/db_xref="taxon:10090"
	/clone="5830469C09"
	/sex="male"
	/tissue_type="thymus"
	/clone_lib="RIKEN Full-length enriched mouse cDNA library /dev_stage="adult"
	1..3259
	/note="similar to RAB3-GAP REGULATORY DOMAIN [Homo sapiens] (SPTRL O75872, evidence: PASTY, 8%ID, 76.7%length, match=3138)"
misc_feature	1..3259
ORIGIN	
Alignment Scores:	
Pred. No.:	3259
Score:	5.55e+04
Length:	8.00
Percent Similarity:	100.00%
Conservative:	0
Percent Identity:	100.00%
Mismatches:	0
Query Match:	1.20%
DB:	11
Gaps:	0
US-10-029-345A-109 (1-665) x AK030955 (1-3259)	
OY	309 SerLyLeuLyLeuLeuHisLeu 316

Db	983	AGTAATTAAACTACTACTT 1006
RESULT 42		
LOCUS	AK031191	
DEFINITION	AK031191	3335 bp mRNA linear HTC 18-SEP-2003
ACCESSION	AK031191	Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length
VERSION	AK031191.1	enriched library, clone:5530426E6 product:similar to RMB3-CAP
KEYWORDS	REGULATORY DOMAIN [Homo sapiens], full insert sequence.	
SOURCE	AK031191	
ORGANISM	AK031191.1 GI:26327104	
REFERENCE	HTC; CAP trapper.	
AUTHORS	Mus musculus (house mouse)	
TITLE	Mus musculus	
JOURNAL	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
MEDLINE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
PUBMED	99279253	
PUBMED	10349636	
REFERENCE	2	
AUTHORS	Carninci, P. and Hayashizaki, Y.	
TITLE	1	
JOURNAL	High-efficiency full-length cDNA cloning	
MEDLINE	Wetli. Enzymol. 503, 19-44 (1999)	
PUBMED	99279253	
PUBMED	10349636	
REFERENCE	2	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,	
TITLE	3	
JOURNAL	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,	
MEDLINE	Komno, H., Akiyama, J., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M.,	
PUBMED	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,	
PUBMED	Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasai, M.,	
REFERENCE	Fujisawa, S., Inoue, K., Togawa, Y., Irawa, M., Ohara, E., Matsubiki, M.,	
AUTHORS	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,	
TITLE	4	
JOURNAL	Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.	
MEDLINE	RIKEN integrated sequence analysis (RISA) system--384-format	
PUBMED	sequencing pipeline with 384 multicapillary sequencer	
PUBMED	Genome Res. 10 (11), 1757-1771 (2000)	
REFERENCE	11076861	
AUTHORS	4	
TITLE	THE RIKEN Genome Exploration Research Group Phase II Team and the	
JOURNAL	PANTOM Consortium.	
MEDLINE	Functional annotation of a full-length mouse cDNA collection	
PUBMED	Nature 409, 685-690 (2001)	
PUBMED	11076861	
REFERENCE	5	
AUTHORS	THE RIKEN Genome Exploration Research Group Phase II Team and the	
TITLE	Group Phase I & II Team.	
JOURNAL	Analysis of the mouse transcriptome based on functional annotation	
MEDLINE	of 60,770 full-length cDNAs	
PUBMED	Nature 420, 563-573 (2002)	
PUBMED	6 (bases 1 to 3335)	
REFERENCE	Adachi, J., Aizawa, K., Akiyama, T., Arai, K., Bono, H., Carninci, P.,	
AUTHORS	Fukuda, S., Furuta, M., Hanagaki, T., Hara, A., Hashizume, W.,	
TITLE	Hayashizaki, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,	
JOURNAL	Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,	
MEDLINE	Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M.,	
PUBMED	Koyu, S., Kurihara, C., Matsuyama, T., Miyazaki, R., Murata, M.,	
PUBMED	Nakamura, Y., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohtsuo, N.,	
REFERENCE	Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,	
AUTHORS	Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,	
TITLE	Sugahara, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akahira, S.,	
JOURNAL	Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,	
MEDLINE	Muramatsu, M. and Hayashizaki, Y.	
PUBMED	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of	
PUBMED	Physical and Chemical Research (RIKEN), Laboratory for Genome	
REFERENCE	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),	
AUTHORS	RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,	
TITLE	kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp).	

COMMENT

URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES

source

Location/Qualifiers

CDS

```

1..3335
  organism="Mus musculus"
  mol_type="mRNA"
  strain="C57BL/6J"
  db_xref="PANTOM DB:5930426E16"
  db_xref="WGI:2354389"
  db_xref="taxon:10090"
  clone="5930426E16"
  tissue_type="forelimb"
  dev_stage="13 days embryo"
  /dev_stage="13 days embryo"
  /clone_lib="RIKEN full-length enriched mouse cDNA library"
  142..3147
  /note="unnamed protein product; putative
  similar to RAB3-GAP REGULATORY DOMAIN [Homo sapiens]
  (SPTT:075872, evidence: FASTY, 87%ID, 76.7%length,
  match=3138)"
  /codon_start=1
  /protein_id="BAC27296.1"
  /db_xref="GI:26327105"
  /translation="MEPATPLAVRFGILPDSRRHGESICLSPCNTLAATVDDEGRVILL
  DVARGIAIRMKGRDQVGMIVQIEDLHRRVVEKGFSPFNTQPSRVAQFLVYA
  PRGIEVNWVQOQPRVGNVNGHGLYPGYKIMGNVNTQSQSPOTQICLVDP
  VSAQVKNVVPFLALSDKSKERAKDLHLKALLRKSPSPDEFEIKELIDI
  KYPATKQALLESILASDRLSFSCIRNVOTIMLTILKNOLSEVDEBLQPCSKATL
  HLYESVQALNTLIDHSPDPSNDLAVLRIDKELIKRALKERIKQNTATVRS
  EDADRVLPVTPLELYEYECDAISIRKVEBECVALGSFPWCLHGSSSTEMCHSL
  ESAGLSFDQLSLLSLWSKEDILDKPQSVCLHTMTLSLKMVAIDETWDSQSV
  SPWMOQRMAICIQSENSGALLSAHGVSAQMSGADKFSQWVADADALDLS
  WEALSDTEYWKLLRLQLEDCILITLHSKLSPPAKAPLSQSEPLDLSVKLIKLG
  GKGIADSVAKWIFKODLSPELKCANKERDVNPEPEDILHLAVRPPSILSDY
  TLMDKVGSKPDRICRDRGMSDTLITLFSGLCELTLSLEADISREVOVPVLT
  EDWLVSEBGISIVELALRQKPTHPVLSHVSLSITLVSARFLSKSVKPLADPSK
  GKNAFKDLITSLIQLPSGEMDNFISVRQOFLKVVSAVQASHKQDPSSEANLTH
  WKDLNPGIADVLAHLIQVSDVIRRHVYGEIYSGADLGEALFOVDKREVLASOL
  LVITGQRLAHLPHOTKEGMELIARLPITLCTWIKAMNPOLQNTGVPDIATATLVH
  KMWELPERKHGGYSLAHLIDVNEANATL"

```

ORIGIN

Alignment Scores:

Pred. No.: 5.69e+04 Length: 3335
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x AK031191 (1-3335)

QY 309 SerTysLeuTylLeuLeuHsiLeu 316

DB 952 AGTAATTAACTACTACATCTT 975

RESULT 43

AK029006/c

LOCUS

DEFINITION

AK029006 3610 bp mRNA linear HTC 18-SEP-2003
 Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
 library, clone:4732460N13 product:dnad (Hsp40) homolog, subfamily
 C, member 5, full insert sequence.

ACCESSION

AK029006 AK029006.1 GI:26324963

KEYWORDS

HTC; CAP trapper.

SOURCE

ORGANISM

Mus musculus (house mouse)

Mus

musculus

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

JOURNAL

High-efficiency full-length cDNA cloning

MEDLINE

99279253

PUBMED

10349636

AUTHORS

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

TITLE

3 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

JOURNAL

Normalization and subtraction of cap-trapper-selected cDNAs to

MEDLINE

2049374

PUBMED

11042159

AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

TITLE

4 Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,

JOURNAL

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,

MEDLINE

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasai, K.,

PUBMED

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,

AUTHORS

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawat, J.,

TITLE

5 Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

JOURNAL

RIKEN integrated sequence analysis (RISA) system-384-format

MEDLINE

Genome Res. 10 (11), 1757-1771 (2000)

PUBMED

20530913

AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the

TITLE

5 PANTOM Consortium.

JOURNAL

Functional annotation of a full-length mouse cDNA collection

MEDLINE

Nature 409, 685-690 (2001)

PUBMED

11076861

AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the

TITLE

5 PANTOM Consortium.

JOURNAL

Functional annotation of a full-length mouse cDNA collection

MEDLINE

Nature 409, 685-690 (2001)

PUBMED

11076861

AUTHORS

5 The RIKEN Genome Exploration Research Group Phase II Team and the

TITLE

6 PANTOM Consortium.

JOURNAL

Analysis of the mouse transcriptome based on functional annotation

MEDLINE

Nature 420, 563-573 (2002)

PUBMED

6 (bases 1 to 3610)

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Aikawa, T., Bono, H., Carninci, P.,

TITLE

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

JOURNAL

Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozawa, T.,

MEDLINE

Hoti, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

PUBMED

Katoch, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

AUTHORS

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

TITLE

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,

JOURNAL

Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

MEDLINE

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

PUBMED

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

AUTHORS

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,

TITLE

Muramatsu, M. and Hayashizaki, Y.

JOURNAL

Direct Submission

MEDLINE

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

PUBMED

Physical and Chemical Research (RIKEN), Laboratory for Genome

AUTHORS

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

TITLE

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

JOURNAL

Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,

MEDLINE

URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,

PUBMED

Fax: 81-45-503-9216)

AUTHORS

cDNA library was prepared and sequenced in Mouse Genome

TITLE

Encyclopedia Project of Genome Exploration Research Group in Riken

JOURNAL

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

MEDLINE

Division of Experimental Animal Research in Riken contributed to

PUBMED

prepare mouse tissues.

AUTHORS

Please visit our web site for further details.

TITLE

URL: <http://genome.gsc.riken.go.jp/>

JOURNAL

URL: <http://fantom.gsc.riken.go.jp/>.

MEDLINE

Location/Qualifiers

PUBMED

1..3610

/organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:4732480N13"
 /db_xref="MGI:2391035"
 /db_xref="taxon:10090"
 /clone="4732480N13"
 /tissue_type="skin"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="10 days neonate"
 163..759
 /note="unnamed protein product; DnaJ (Hsp40) homolog,
 subfamily C, member 5 (MGDI[MGI:892995, GI|NM_016775,
 evidence: BLASTN, 99%, match=557])
 putative"
 /codon_start=1
 /protein_id="BAC26236.1"
 /db_xref="GI:26324964"
 /translation="MADRORSLSITSGESLHYVLGDKNATSDDIKKSRYRLAKYHPR
 DKNDPNPADKFEKINNNAHILDTATRNIVYDKGSLGLYVAQFGENNTYFVLS
 SWMAKLFVVGGLTCYCCCLCCFCNCCGCKPKPABEEETFPVSPEDLAQLQ
 SDERADTPIVLIQPASATETTLTADSHPSYHTDGFN"

ORIGIN
 Alignment Scores:
 Prod. No.: 6.19e+04 Length: 3610
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.20% Indels: 0
 DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x AK029006 (1-3610)
 Oy 95 ValAlaSerLeuSerSerAspCys 102
 |||||
 Db 662 GTAGCCTCCCTTTCATCAGACTGC 639

RESULT 44
 AK046988
 LOCUS
 DEFINITION
 Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length
 enriched library, clone:B930009J12 product:unclassified, full
 insert sequence.
 AK046988
 AK046988.1 GI:26091852
 HTG; CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

3617 bp mRNA linear HTC 20-SEP-2003
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 2 99279253
 10349636
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Komoto, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuzaki, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

CDS
 REFERENCE
 JOURNAL
 MEDLINE
 PUBMED
 PUBLISHED
 11042359
 11042359
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 PUBLISHED
 11042359
 11042359
 AUTHORS

TITLE	RIKEN integrated sequence analysis (RISA) system--344-format
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	
AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 3617)
AUTHORS	Adachi, J., Aizawa, K., Akamizu, T., Arakawa, T., Bono, H., Carrincci, P., Fukuda, S., Furuno, M., Hamaguchi, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiwamoto, K., Hiraoka, T., Hirozane, T., Horii, F., Imclanti, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, url: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/.
FEATURES	Location/Qualifiers 1. 3617 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:B930009J12" /db_xref="MGI:2412520" /db_xref="taxon:10090" /clone="B930009J12" /tissue_type="cerebellum" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="10 days neonate" 1. 3617 /note="unclassifiable"
ORIGIN	misc_feature
Alignment Scores:	6.21e+04 Length: 3617
Pred. No.:	8.00 Matches: 8
Score:	100.00% Conservatve: 0
Percent Similarity:	100.00% Mismatches: 0
Best Local Similarity:	1.20% Indels: 0
Query Match:	11 Gaps: 0
DB:	
US-10-029-345A-109 (1-665) x AK046988 (1-3617)	
xy	163 LeuProbAnleuTyfLeuGlyCyG 170

Db 3401 TTACCAACTGTATGTAGATGT 3424

RESULT 45
AK032373/c

LOCUS AK032373 3626 bp mRNA linear HTC 18-SEP-2003

DEFINITION Mus musculus adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:6430529B06 product:DnaJ (Hsp40) homolog, subfamily C, member 5, full insert sequence.

ACCESSION AK032373

VERSION AK032373.1 GI:26328200

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Smit, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M., Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Yoneda, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

MEDLINE 11076861

PUBMED 11076861

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 11076861

PUBMED 11076861

REFERENCE 6

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komoto, H., Konda, M., Koyata, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toy, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-Jul-2001) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

COMMENT

Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers

1.3626

source

FEATURES

source

1.3626

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM DB:6430529B06"

/db_xref="MG1:2395854"

/db_xref="taxon:10090"

/clone="6430529B06"

/sex="male"

/tissue_type="olfactory brain"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

157..753

/note="unnamed protein product, DnaJ (Hsp40) homolog, subfamily C, member 5 (MG1:892995, GB|NM_016775, evidence: BLASTN, 99%, match=597)

putative"

/codon_start=1

/protein_id="BAC27841.1"

/db_xref="GI:26328201"

/translation="MAPQRORSSTSGESLYHVLGDKNATSDDIKSYKALKATPH DKRDNPEADKDEINNAHILTDATKRNITDYKSLGLVVAQFEENVVNYFVLS SWMAKALFVCGGLTCCYCCCLCCCFCCGCKPKAPGEGTEFFVSPEDLEAQLQ SDEAREADPTVLPASATETQTLDSDHSPHYTDGFN"

ORIGIN

Alignment Scores:

Pred. No.: 6.22e+04 Length: 3626

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.20% Indels: 0

DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x AK032373 (1-3626)

Oy 95 ValAlaSerLeuSerSerAspCys 102

Db 656 GTAGCTTCCTTCATCAGACTGC 633

RESULT 46

LOCUS AK083783 3701 bp mRNA linear HTC 20-SEP-2003

DEFINITION Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone: D130008F08 product: unclassified, full insert sequence.

ACCESSION AK083783

VERSION AK083783.1 GI:26101489

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to

JOURNAL MEDLINE PUBMED 20499374 11042159
REFERENCE AUTHORS
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,Y., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE PUBMED 20530913 11076861
REFERENCE AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL MEDLINE PUBMED 20499374 11042159
REFERENCE AUTHORS
5
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL MEDLINE PUBMED 20499374 11042159
REFERENCE AUTHORS
6
(bases 1 to 3701)
Adachi,J., Aizawa,K., Akiyama,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imocant,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohnato,N., Okazaki,Y., Saito,R., Satoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shingawa,K., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

JOURNAL MEDLINE PUBMED 20530913 11076861
REFERENCE AUTHORS
7
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. 3701
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:D130008F08"
/db_xref="WGI:2419339"
/db_xref="taxon:10090"
/clone="D130008F08"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
1. 3701
/note="unclassifiedable"

JOURNAL MEDLINE PUBMED 20499374 11042159
REFERENCE AUTHORS
8
misc_feature
1. 3701
/note="unclassifiedable"

JOURNAL MEDLINE PUBMED 20499374 11042159
REFERENCE AUTHORS
9
misc_feature
1. 3701
/note="unclassifiedable"

Alignment Scores:
Pred. No.: 6.36e+04 Length: 3701
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 11 Gaps: 0

US-10-029-345a-109 (1-665) x AK083783 (1-3701)
QY 359 SerValProSerValProSerVal 366
|||||
Db 2227 TCTGTGCCCTCTGTGCCCTCTGTG 2250
|||||

RESULT 47
AK084916 3722 bp mRNA linear HTC 20-SEP-2003
AK084916
LOCUS
DEFINITION
Mus musculus 13 days embryo lung cDNA, RIKEN full-length enriched library, clone:ID430011E20 product:hypothetical protein, full insert sequence.
ACCESSION
AK084916
VERSION
AK084916.1 GI:26351342
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
1
Carninci,P. and Hayashizaki,Y.
TITL
HIGH-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL MEDLINE PUBMED 10349636
REFERENCE AUTHORS
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE PUBMED 11042159
REFERENCE AUTHORS
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,Y., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE PUBMED 20530913 11076861
REFERENCE AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL MEDLINE PUBMED 20499374 11042159
REFERENCE AUTHORS
5
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL MEDLINE PUBMED 20499374 11042159
REFERENCE AUTHORS
6
(bases 1 to 3722)
Adachi,J., Aizawa,K., Akiyama,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imocant,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohnato,N.,

TITLE
JOURNAL

Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

COMMENT

Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Tsurumi-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopaedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
Location/Qualifiers

FEATURES

source

CDS

1..3722
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:D430011E20"
/db_xref="MGJ:2421716"
/db_xref="taxon:10090"
/clone="D430011E20"
/tissue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
992..1294
/note="unnamed protein product; hypothetical protein (evidence: decoder, longest-ORF)
putative"
/codon_start=1
/protein_id="BAC39308.1"
/db_xref="MGJ:26351343"
/translation="MEQITDPAVLVQVITLMLSHLSPDMFSGVSGRLVSALQLSMQRT
WPPFVGPVIAFLHLNKNRKLFGGLCPQLVLYLHRSWAFSASRPVGDCE"

ORIGIN

Alignment Scores:

Pred. No.: 6.4e+04 Length: 3722
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x AK084916 (1-3722)

QY 357 ProAlaSerValProSerValPro 364

DB 1207 CCAGCGTGTGTACTTCTGTACTT 1230

RESULT 48

BC049163
LOCUS BC049163 3722 bp mRNA linear HTC 17-DEC-2003
DEFINITION Mus musculus cDNA clone IMAGE:6466814, containing frame-shift errors.
ACCESSION BC049163
VERSION BC049163.1 GI:29165699
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE

1 (bases 1 to 3722)
Strusberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

REMARK
COMMENT

Altechul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stachewicz, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Ueda, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loguercio, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.J., Walek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huliyil, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeslee, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Kitzwinski, M.I., Skalka, U., Smalins, D.E., Schercher, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3722)
Strusberg, R.
Direct Submission
Submitted (21-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (ILNL)
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahy, Erin Helton, Mark Ketterman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAP Plate: 105 Row: 1 Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27369777
This clone has the following problem: frame shifted.

FEATURES

source

1..3722
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6466814"
/tissue_type="Eye, retina, mouse strain C57BL/6"
/clone_lib="NIH MGC_94"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:

Pred. No.: 6.4e+04 Length: 3722
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 11 Gaps: 0

US-10-029-345A-109 (1-665) x BC049163 (1-3722)

QY 410 SerAlaSerMetAlaAlaSerLeu 417

DB 2435 AGTCATCATGAGCAGCACTT 2458

RESULT 49

AK035810

LOCUS AK035810 3727 bp mRNA linear HTC 19-SEP-2003
DEFINITION Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:9630006820 product:POLY(A) POLYMERASE GAMMA
(EC 2.7.7.19) (NEO-POLY(A) POLYMERASE) homolog [Homo sapiens], full
insert sequence.
ACCESSION AK035810
VERSION AK035810.1 GI:26330927
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PubMed 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PubMed 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
Komno, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, S., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Iwasa, M., Ogata, E., Matsuoka, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, D.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PubMed 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Aichi, J., Aizawa, K., Akimura, T., Arai, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirao, T., Hirazawa, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kankawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Koda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ono, M., Ono, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
Yamamoto, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>
location/Qualifiers

FEATURES
source
1.3727
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9630006820"
/db_xref="MGI:2400157"
/db_xref="taxon:10090"
/clone="9630006820"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
216..2435
/note="unnamed protein product; POLY(A) POLYMERASE GAMMA
(EC 2.7.7.19) (NEO-POLY(A) POLYMERASE) homolog [Homo
sapiens] (SPR|Q969N1, evidence: FASTA, 93.5%ID,
100%length, match=2217)
putative"
/codon_start=1
/protein_id="BAC29194.1"
/db_xref="GI:26330928"
/translation="MKMSGANTMLDSQROOKHYGTSPISLACPKEDIDHYTKLIDA
MKPFGVEEBEELNHLVYLKGNLNVKMSIDISSEKSLPSPVAVTGGKIPFSY
RGVHTKGADIDLCVAPRHSRDFOSFEKRLKQDGRNLRVAVDAEPVIRKEF
DGIEITLVARALRIOTISNLDRLSDRLSDIRLSINGCRVADILHYPNKT
FRITLRAVYLAKRRTGISYNMGLFQVSMHMLVATCOLYPRMAASTLVKRFVPS
KWEKPVVLKQPEESNLNLVWDVNSDRHMLPITPAIPQONSTYVNSTRT
VWVEKQKQIAVDELILQKSDMSKLEPTNPFQKRYHIVLASASTENHLEWGL
VSKIRVLGNLENEFILTAVNPOSFGNKHANNVSMFLGILFRVYKAS
VNIDLVDIOSFTDYVROANNINMLKQMKIETHVKKKHLHYLPAILQKKSL
SPVSRSGGLQSRSLSDTCLDSRDTSCIPPNPVSANRNSDPSPTGTERESA
EPVAVYVEKLPSVPPAQGISPTVAKVDPTAKAVSPAVCTIPYVGNVIRITTP
HNPOQPHLNGISNITKVTNTPRKRSLPPTDGSRLKQIKERILRESAFKESRAED
RRKPMDSIGSMPITPTIDARKRLPSKELEPDSPPVANNIRVINKSIRLTNR"

CDS
6
/note="unnamed protein product; POLY(A) POLYMERASE GAMMA
(EC 2.7.7.19) (NEO-POLY(A) POLYMERASE) homolog [Homo
sapiens] (SPR|Q969N1, evidence: FASTA, 93.5%ID,
100%length, match=2217)
putative"
/codon_start=1
/protein_id="BAC29194.1"
/db_xref="GI:26330928"
/translation="MKMSGANTMLDSQROOKHYGTSPISLACPKEDIDHYTKLIDA
MKPFGVEEBEELNHLVYLKGNLNVKMSIDISSEKSLPSPVAVTGGKIPFSY
RGVHTKGADIDLCVAPRHSRDFOSFEKRLKQDGRNLRVAVDAEPVIRKEF
DGIEITLVARALRIOTISNLDRLSDRLSDIRLSINGCRVADILHYPNKT
FRITLRAVYLAKRRTGISYNMGLFQVSMHMLVATCOLYPRMAASTLVKRFVPS
KWEKPVVLKQPEESNLNLVWDVNSDRHMLPITPAIPQONSTYVNSTRT
VWVEKQKQIAVDELILQKSDMSKLEPTNPFQKRYHIVLASASTENHLEWGL
VSKIRVLGNLENEFILTAVNPOSFGNKHANNVSMFLGILFRVYKAS
VNIDLVDIOSFTDYVROANNINMLKQMKIETHVKKKHLHYLPAILQKKSL
SPVSRSGGLQSRSLSDTCLDSRDTSCIPPNPVSANRNSDPSPTGTERESA
EPVAVYVEKLPSVPPAQGISPTVAKVDPTAKAVSPAVCTIPYVGNVIRITTP
HNPOQPHLNGISNITKVTNTPRKRSLPPTDGSRLKQIKERILRESAFKESRAED
RRKPMDSIGSMPITPTIDARKRLPSKELEPDSPPVANNIRVINKSIRLTNR"

ORIGIN
Alignment Scores:
Pred. No.: 6,41e+04 Length: 3727
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 11 Gaps: 0
US-10-029-345A-109 (1-665) x AK035810 (1-3727)
Qy 410 Ser1aseMera1a1ase1eu 417
Db 2474 AGTCATCGATGACACCACTCTT 2497
RESULT 50
ACCESSION AK078517 3789 bp mRNA linear HTC 18-SEP-2003
VERSION AK078517
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE

AUTHORS

TITLE

JOURNAL Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 MEDLINE Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 PUBMED Normalization and subtraction of cap-trapper-selected cDNAs to
 20499374 prepare full-length cDNA libraries for rapid discovery of new genes
 11042159 Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE

AUTHORS

TITLE

JOURNAL Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 MEDLINE Kono, H., Akiyama, J., Nishi, K., Kitunai, T., Tashiro, H., Itoh, M.,
 PUBMED Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 11076861 Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashwagi, K.,
 4 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Ohara, E., Watahiki, M.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

REFERENCE

AUTHORS

TITLE

JOURNAL The RIKEN Genome Exploration Research Group Phase II Team and the
 MEDLINE FANTOM Consortium.
 PUBMED Functional annotation of a full-length mouse cDNA collection
 5 Nature 409, 685-690 (2001)

REFERENCE

AUTHORS

TITLE

JOURNAL The FANTOM Consortium and the RIKEN Genome Exploration Research
 MEDLINE Group Phase I & II Team.
 PUBMED Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 6 Nature 420, 563-573 (2002)

REFERENCE

AUTHORS

TITLE

JOURNAL Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 MEDLINE Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 PUBMED Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 11076861 Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, C., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M., and Hayashizaki, Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

COMMENT

CDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers

FEATURES

source

1. 3789
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6J"
 /db_xref="FANTOM_DB:682042BL09"
 /db_xref="MGI:2396740"
 /db_xref="taxon:10090"
 /clone="682042BL09"

misc_feature
 1. 3789
 /note="weakly similar to KIAA1614 PROTEIN (FRAGMENT) (Homo
 sapiens) (SPT1|Q9HCF8, evidence: FASTV, 57.3%ID,
 68.7%length, match=2504)"

ORIGIN

Alignment Scores:
 Pred. No.: 6.52e+04
 Score: 8.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.20%
 DB: 11
 Gaps: 0

US-10-029-345A-109 (1-665) x AK078517 (1-3789)

QY 544 LeuAlaProGlnThrSerThrPro 551

Db 2171 CTGGCTCCACAGACGACGACCC 2194

Search completed: June 22, 2004, 10:09:34
 Job time : 437 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2004, 03:17:17 ; Search time 664 Seconds
(without alignments)
4254.596 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665
Sequence: 1 MAHEMIGCTIVTERLVALL.....LKVSGSSFSGSMETIEVS 665

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 212409041 residues

Word size: 1

Total number of hits satisfying chosen parameters: 772722

Minimum DB seq length: 1419

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q/cg2.1/USPTO.spool/US10029345/runat.21062004.164401.29745/app.query.fasta.1.839
-DB=N Geneseq.29Jan04 -QPMT=fastlap -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=75 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEA=SIZE=500 -MINLEN=1419 -MAXLEN=200000000
-USER=US10029345.@CGN_1_1.470.@runat.21062004.164401.29745 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	665	100.0	5450	6	ACCG60559
2	665	100.0	5450	6	ACCG60572
3	472	71.0	2132	7	ADA53105
4	472	71.0	2732	4	AAD09492
5	472	71.0	2966	4	AAH99685
6	472	71.0	3059	4	AAH99685
7	472	71.0	3104	6	ABN59704
8	472	71.0	3332	6	ABK48378

9	472	71.0	3496	6	ABK47596	Abk47596 cDNA enco
10	472	71.0	3544	6	AA514639	AA514639 Human cDN
11	472	71.0	3544	6	ABK49402	Abk49402 cDNA enco
12	472	71.0	3766	6	ABK14474	Abk14474 Human pro
13	472	71.0	4790	6	ABN83966	Abn83966 Human gen
14	472	71.0	5145	5	ABV20833	Abv20833 Human pro
15	472	71.0	5145	5	ABV21080	Abv21080 Human pro
16	472	71.0	5145	5	ABV26680	Abv26680 Human pro
17	472	71.0	5145	5	ABV20978	Abv20978 Human pro
18	472	71.0	5145	5	ABV21092	Abv21092 Human pro
19	472	71.0	5145	5	ABV21312	Abv21312 Human pro
20	472	71.0	5145	5	ABV21316	Abv21316 Human pro
21	472	71.0	5145	5	ABV26826	Abv26826 Human pro
22	472	71.0	5145	5	ABV27131	Abv27131 Human pro
23	472	71.0	5145	5	ABV26923	Abv26923 Human pro
24	472	71.0	5145	5	ABV27135	Abv27135 Human pro
25	472	71.0	5145	5	ABV28657	Abv28657 Human pro
26	472	71.0	5145	5	ABV22827	Abv22827 Human pro
27	472	71.0	5145	5	ABV26934	Abv26934 Human pro
28	472	71.0	5145	5	AAE30479	AAE30479 Human pro
29	472	71.0	5111	6	ACCG60521	ACCG60521 Polynucle
30	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
31	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
32	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
33	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
34	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
35	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
36	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
37	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
38	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
39	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
40	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
41	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
42	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
43	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
44	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
45	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
46	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
47	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
48	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
49	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
50	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
51	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
52	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
53	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
54	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
55	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
56	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
57	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
58	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
59	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
60	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
61	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
62	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
63	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
64	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
65	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
66	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
67	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
68	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
69	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
70	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
71	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
72	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
73	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
74	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle
75	472	71.0	5111	6	ACCG60560	ACCG60560 Polynucle

RESULT 1
ACCG60559

ALIGNMENTS

ID ACC60559 standard; cDNA; 5450 BP.
 AC ACC60559;
 XX
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Polynucleotide relating to the invention SEQ ID NO: 108.
 XX
 KW Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic;
 KW antiproliferative; cardiac; cytostatic; gene therapy; liver disease;
 KW proliferative disorder; renal failure; cardiovascular disorder;
 KW immunological disorder; arthritis; psoriasis; congenital heart defect;
 KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
 XX
 OS Homo sapiens.
 XX
 PN W0200257460-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 20-DEC-2001; 2001WO-US050459.
 XX
 PR 20-DEC-2000; 2000US-0256868P.
 PR 30-MAR-2001; 2001US-0280186P.
 PR 01-MAY-2001; 2001US-0287735P.
 PR 05-JUN-2001; 2001US-0295848P.
 PR 25-JUN-2001; 2001US-0300465P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
 PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
 PI Kyresk S, Mcatee P, Suchard S, Banas D;
 XX
 DR WPI: 2002-599721/64.
 DR P-PSDB; ABR52381.
 XX
 PT Novel polynucleotides encoding human phosphatase polypeptides useful in
 PT the prevention or treatment of e.g. proliferative and cardiovascular
 PT disorders.
 XX
 PS Claim 1; Fig 13; 801BP; English.
 XX
 CC The invention relates to a novel isolated nucleic acid comprising a
 CC polynucleotide having a nucleotide sequence selected from 40
 CC polynucleotides fully defined in the specification. The polynucleotide of
 CC the invention has antiproliferative, hepatotropic, nephrotropic,
 CC antitumor, antiproliferative, cardiac, and cytostatic activity. The
 CC polynucleotide may have a use in gene therapy. A polynucleotide or
 CC polypeptide of the invention is useful for preventing, treating or
 CC ameliorating a medical condition, e.g. a proliferative disorder. They are
 CC also useful for treating e.g. liver disease, renal failure, immunological
 CC disorders including arthritis and psoriasis, cardiovascular disorders
 CC such as congenital heart defects and congestive heart failure, and
 CC cancer. A method of the invention is useful for diagnosing a pathological
 CC condition or susceptibility to a condition in a subject. The present
 CC sequence is used in the exemplification of the invention
 XX
 SQ Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 5450
 Score: 665.00 Matches: 665
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-345A-109 (1-665) x ACC60559 (1-5450)
 OY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 Db 538 ATGGCCCATGATGATATGGAACTCAAAATTGTACTGAGAGGTGTGGCTCTGCTGGAA 597

OY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
 Db 598 AGTGGAAACGGAAGAAAGTGTCTTAATGTATGATCCGGCATTTGTGAATATCAATCATCC 657
 OY 41 HisIleLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
 Db 658 CACATTTTGGAAAGCCATTATATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 717
 OY 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
 Db 718 GACAAAGCTTATATACAGAGCTCATCAAGCATTTACAGGAACATTAAGCTTACATTAAT 777
 OY 81 CysSerGluValValValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 Db 778 TGCAGTCAAGAGGTGTGATTATCGATCAAGCTCCAAAGTGTGCTCTCTCTCTCA 837
 OY 101 AspCysPheLeuThrValLeuLeuGlyIleLeuGlyIleValSerPheAsnSerValHisLeu 120
 Db 838 GACTGTCTTCTACTGTAATCTCTGGTAAACCTGAGAAAGACCTTCACCTGTCACCTG 897
 OY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGlyGlyIleSerSer 140
 Db 898 CTGGCAGGTGGGTTTGTGAGATTCTCTGCTGTTTCCCTGGCCTCTGGAAGAAATCC 957
 OY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 Db 958 ACTTAGTCCCTACCTGATTTCTCAGCTTCTTACCTGTTTACCTGTTTACCTGTTTAC 1017
 OY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnValGluLeuIle 180
 Db 1018 CGAATTTCTCCCAATCTTTATCTTGCTGCTGCGCAGAGAGTGTCTCAACAGAGACTGATA 1077
 OY 181 GlnGlnAsnGlyIleGlyTyrValIleAsnAlaSerTyrThrCysProIleProAspPhe 200
 Db 1078 CAGCAGATGGGATGGTATGTATGTAAATCCAGCATATACCTGTCCAAAGCTGACCTT 1137
 OY 201 IleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlyIleLeu 220
 Db 1138 ATCCCGAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1197
 OY 221 ProTyrLeuAspIleValSerValAspPheIleGluValAlaValAsnArgGlyCysVal 240
 Db 1198 CCGTGTGGTGGCAATATGATGATTTTCAATTCAGAAAGCAAAAGCTCAATGATGATG 1257
 OY 241 LeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaMet 260
 Db 1258 CTAGTGCATGTGTAGCTGGGATCTCCGCTCCGCAACATCGCTATCGCTATCATCATG 1317
 OY 261 LysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGlyValArgProThr 280
 Db 1318 AAGAGATGAGCATGTCTTATGATGAGAGCTTACGATTTGTGAAGAAAGAAAGCTCACT 1377
 OY 281 IleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlnValIleValAsn 300
 Db 1378 ATATCTCAAACTTCATTTTCTGGGCAACTCTGTGACATATGAGAAAGATTAAGAAC 1437
 OY 301 GlnThrGlyAlaSerGlyProIleSerIleLeuValIleLeuValIleLeuValIleLeu 320
 Db 1438 CAGACTGAGCATCAGGCGCAAGAGCAACTCAAGCTGTGCACTCGAGAGGCCAAAT 1497
 OY 321 GluProValProAlaValIleSerGlyGlyGlnIleValSerGlnThrProLeuSerProPro 340
 Db 1498 GAACTTCTCCCTGTGTCTCAAGAGGGTGGCAAGAAAGCAAGAGCGCCCTCAATCCACC 1557
 OY 341 CysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerVal 360
 Db 1558 TGTGCCACTCTGCTACTCAGAGGAGCAGAGCAAGAGGCCCTGTGATCCCGCAGGCTG 1617
 OY 361 ProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeu 380
 Db 1618 CCCAGCTGCGCCAGCGTGCAGCGCTGCTGTTAGAGAGCAGCGCGCTGTGTAAGCGGCTC 1677

QY	381	SerGlyLeuNH ₂ LeuSerAlaAspArgGluGluAspSerAsnLysLeuValArgSerPhe	400
Db	1678	AGTGGCGCTGCACCTGTCCCGACGACGCTGGAGACACAAATTAAGCTTAACCGTTCCTTC	1737
QY	401	SerLeuAspIleLeuSerValSerTyrSerAlaSerMetAlaAspLeuNH ₂ GlyPhe	420
Db	1738	TCTCGATATCAAAATCAGTTTTCATATTACGACGACATGGCAGCATCTTTCATGGCTTC	1797
QY	421	SerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAsn	440
Db	1798	TCCCTCATCAGAAAGTCTTTGGAAATACACAAACCTTCCACTACTCTGGATGGGACCAAC	1857
QY	441	LysLeuCybGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerPro	460
Db	1858	AAGCTATGCCAGTTCTCCCTGTTCCAGAACATATCGAGACAGACTCCGGAACCAAGTCT	1917
QY	461	AspLysGlnGluAlaSerIleProLysLeuGlnThrAlaArgProSerAspSerGln	480
Db	1918	GATAAGAGGAGAAGCCAGCATCCCGACAGAGCTGCAGACCGCAGCCCTTCAGACAGCCAG	1977
QY	481	SerLysValGluNH ₂ LeuSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu	500
Db	1978	AGCAAGCCATTTGCATTCCGTCAGAAACCGACAGAGTGGCACCGCCACAGAGCTCCCTTTTA	2037
QY	501	SerProLeuNH ₂ ArgSerGlySerTyrAlaGluAspAsnTyrHisThrSerPheLeuPheGly	520
Db	2038	TCTCCACATGCATCGAAGTGGAGCGCTGGAGACAAATTACACACACGCTTCTTTTCGGC	2097
QY	521	LeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPheHis	540
Db	2098	CTTTCACACACCGACGACGACCTCAGCAAGTCTGCTGGCTTGGGCTTTAAGGCTGGGCAC	2157
QY	541	SerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrTyrPheAla	560
Db	2158	TCGGATATCTTGGCCCCCGACACCTCTACCCCTTCCGACACAGACAGCTGGTATTTGGC	2217
QY	561	ThrGluSerSerHisPheTyrSerSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrSer	580
Db	2218	ACAGAGTCTCTACACTTACTCTGCTCCACACCACTTACGAGGACGACGACAGTTAACTCT	2277
QY	581	AlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgArg	600
Db	2278	GCTTACACTGCAGCCAGCTGCTCCACTTGGCGGAGCCAAAGTCTATTCTGTGCGAGCGG	2337
QY	601	GlnLysProSerAspArgAlaAspSerArgTyrSerTyrHisGluGluSerProPheGlu	620
Db	2338	CAGAAAGCAAGTACAGAGCTGACTCGCGGGGAGCTGCATGAAGAAGCCCTTTGAA	2397
QY	621	LysGlnPheLysValArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsn	640
Db	2398	AAGCAGTTTAAACCGACGAAGCTGCCAAATGGAATTTGGAGAGAGCATCATCTCAGAGAAC	2457
QY	641	ArgSerArgGluGlnLeuGlnLysValGlySerGlnSerSerPheSerGlySerMetGlu	660
Db	2458	AGGTACCGGAGAAAGACTGGGGAAAGTGGGACGTCACTTCTTTCGGGACGACATGGA	2517
QY	661	IleIleGluValSer665	
Db	2518	ATCATTTAGAGTCTCC2532	
RESULT 2			
ACC60572			
ID	ACC60572 standard, cDNA, 5450 BP.		
NC	ACC60572;		
XX	19-JUN-2003 (first entry)		
DT			
XX			
DE	Polynucleotide relating to the invention SEQ ID NO: 147.		
XX			
Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic; antipneumatic; cardiac; cyostatic; gene therapy; liver disease; proliferative disorder; renal failure; cardiovascular disorder;			

KW	immunological disorder; arthritis; proslasia; congenital heart defect;
KV	congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX	
OS	Homo sapiens.
XX	
PN	WO200257460-A2.
XX	
PD	25-JUL-2002.
XX	
PF	20-DEC-2001; 2001WO-US050459.
XX	
PR	20-DEC-2000; 2000US-0256866P.
PR	30-MAR-2001; 2001US-0280186P.
PR	01-MAY-2001; 2001US-0287735P.
PR	05-JUN-2001; 2001US-0295848P.
PR	25-JUN-2001; 2001US-0300465P.
XX	
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.
XX	
PI	Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
PI	Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Basolino D;
PI	Krystek S, Meatee P, Suchard S, Banas D;
XX	
DR	WPI; 2002-599721/64.
XX	
XX	P-PSDB; ABR52407.
PT	
PT	Novel polynucleotides encoding human phosphatase polypeptides useful in
PT	the prevention or treatment of e.g. proliferative and cardiovascular
PT	disorders.
XX	
PS	Example 7, Fig 19, 801pp, English

The invention relates to a novel isolated nucleic acid comprising a polynucleotide having a nucleotide sequence selected from 40 polynucleotides fully defined in the specification. The polynucleotide of the invention has antiproliferative, hepatotropic, nephrotropic, antiarthritic, antipsoaritic, cardiant, and cytostatic activity. The polynucleotide may have a use in gene therapy. A polynucleotide or polypeptide of the invention is useful for preventing, treating or ameliorating a medical condition, e.g. a proliferative disorder. They are also useful for treating e.g. liver disease, renal failure, immunological disorders including arthritis and psoriasis, cardiovascular disorders such as congenital heart defects and congestive heart failure, and cancer. A method of the invention is useful for diagnosing a pathological condition or susceptibility to a condition in a subject. The present sequence is used in the exemplification of the invention.

XX Sequence 5450 BP; 1471 A; 1211 C; 1337 G; 1431 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	5450
Score:	665.00	Matches:	665
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-029-345A-109 (1-665) x ACC60572 (1-5450)

QY 1 MecAlaHsGluMetIleGlyThrGlnIleValIThrGluArgLeuValAlaLeuLeuGlu 20
 Db 538 ATGGCCCATGAGATATTGGAACTCAAAATTGTACTGAGAGCTGGCTCTCGCGAA 597
 QY 21 SerGIYThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 Db 558 AGTGAACCGAAGAAAGCTGCTGTAATTGATTACCGCGCAATTGTGGAATAACATACATCC 657
 QY 41 HisIleLeuGluAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsn 60
 Db 658 CACATTGTGGAAGCATTTATATATCAACTGCTCCAAAGCTTATGAGAGAGGTTGCAACAG 717
 QY 61 AspIlyValLeuIleIleThrGluLeuIleGlnIleSerAlaIleHisIleValAspIleAsp 80

```

Db      718 GACAAAGTGTATTACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTTGACATTGAT 777
QY      81 CYSSEGLILYSVALVALVALYRASPGLINSESEGINASPVALALASERLEUSER 100
Db      778 TCCAGTCAGAAAGTTAGTTTACATCAAAAGCTCCCAAGATGTTCCCTCTCTCTTCA 837
QY      101 AAPCYSPHELEUTHRALLEULEUGLYVLSLEUGILYSESPHELSNSERVALHISLEU 120
Db      838 GACGTTTTCTCACTGACTCTCGGGTAAACTGGAAAGAGCTTCAACTCTGTTCACCTG 897
QY      121 LEUALAGLYGYPHELAGLUPHESERARGYSPHEPROGLYLEUCYSGILUGILYUSSE 140
Db      898 CTTGCAGAGTGGGTTTCTGAGTTCTCTCGTTGTTTCCCTGGCCTCTGTGAGAGAAATCC 957
QY      141 THRLEUVALPROTHRYSILSESEGINPROCYGLEUPROVALALASNTLEGLYPROTHR 160
Db      958 ACTCTAGTCCCTAACCGCATTTCTCAGCCTGTCACTGCTTGCCTCAACATTTGGGCCAAC 1017
QY      161 ARGILILEUPROASNLEUTHRYLEUGLYCYSGINARASPAVALLEUASNLISGLILEU 180
Db      1018 CGAATTTCTTCCAAATCTTATCTTGAGCTGCACGAGATGCTCTCAACAGAGACTGATA 1077
QY      181 GINGINASNGLYILEGLYTRYVALLEUASNALASERTYRTHRCSYSPROLYSPROASPH 200
Db      1078 CAGCAAAATGGGATTGGTTATGTTAAATGCCAGCTATACCTGTCAAAAGCCTGACTTT 1137
QY      201 ILEPROGLUSERHISPHLEUARGVALPROVALASNASPSPHECYSGILULYSILEU 220
Db      1138 ATCCCCGAGTCCATTTCTCGCTGCTGCTGTAAGACAGCTTTGTGAGAAATTTTG 1197
QY      221 PROTPRLEUASPLYSSESERVALASPHLEILEGLULYSALALYSALASERASNGLYCYSE 240
Db      1198 CCGGTGTTGACAAATCAGTAGATTTCAATTGAGAAAGCAAAAGCTCCCAATGATGTGT 1257
QY      241 LEUVALHISCYSEUVALAGLYILESEARGSERVALATHRILEALILEALARYLLEWC 260
Db      1258 CTAGTGCATGTTTACCTGGGATCTCCCGCTCGGCACCATCGCTATCCCTAACATCANG 1317
QY      261 LYSARGMETASPMETSERLEUASPGILUALARYRARGPHEVALILYSILYSGILYSGARG 280
Db      1318 AAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAAAAAAGAACCTACT 1377
QY      281 ILESEPROASNPHASNPHLEUGLYGLNLEULEUASPTYRGLULYSILYSILEVSAN 300
Db      1378 ATRTCTCCAACTTCAATTTCTGGGCCAACTCTGGACTATGAGAAAGATTAAAGAAC 1437
QY      301 GINTHRGILYALASERGLYPROLYSERLYSLEULYSLEULEUHSLEUGILULYSPROAN 320
Db      1438 CAGACTGGAGCATCAGAGGCCAAAGAACCAACTCAAGCTGTGACCTGAGAAAGCCAAAT 1497
QY      321 GIUPROVALPROVALVALSERGLUGLYGLNLYSESGILUTHRPROLEUSERPROPO 340
Db      1498 GAACCTGCTCCCTGCTCTCAGAGGGTGAACAGAAAAAGGAGAGCGCCCTCAGTCCACCC 1557
QY      341 CYSAIASPSEVALATHRSESGILUALAGLYGLNARGPROVALHISPROLASERVAL 360
Db      1558 TGTGCCGACTCTCTACCTCAGAGGACAGACAGCAAAAGCCCTGATCCCGCACCGTG 1617
QY      361 PROSERVALPROSERVALGINPROSERLEULEUGILUASPSERPROLEUVALGINALAEU 380
Db      1618 CCAGAGGTGCCCGCTGACAGCGCTGCTGTAGAGAGACGCCGCTGTGATACAGGCGCTC 1677
QY      381 SERGLYLEUHSILEUSERVALASPARGLUUGILUASPSERASNLYSLEULYSARGSERPHE 400
Db      1678 AGTGGGCTGACCTGTCCGACAGCAGGCTGGAAGACAGCAATTAAGCTCAAGCCTTCTTC 1737
QY      401 SERLEUASPILEYSESERVALSERTYRSEVALASERMETALALASERLEUHSILYSPHE 420
Db      1738 TCTCTGATATCAAAATCAGTTTCAATTCAGCCAGATGCGACATCTTACATGCTTCC 1797
QY      421 SERSESEGLUASPALEUGILYTRYRYSERPROSETHRTHREUASPGILYTHRAAN 440
Db      1798 TCTCATCAGAAAGATGCTTTGGAATACTCAAACTTCCACTGTGATGGAGCAAC 1857

```

```

QY      441 LYSLEUCYSGILNPHESERPROVALGINGLILEUSERGLUGILNTHRPROGLUTHRSEPRO 460
Db      1858 AAGCTATCCAGATTCTCCCTGTGTGAGAACTATCGAGACAGATCCCGAAACAGTCT 1917
QY      461 AAPLYSGILUALASERILEPROLYSILEUGILNTHRALARGPROSERASPSERGIN 480
Db      1918 GATAAGAGGAAGACAGCATCCCAAGAAAGCTGACAGACCGCAGGCTTTCAGACAGCCAG 1977
QY      481 SERLYSARGLEUHSISERVALARGTHRISERSESERGLYTHRALAGLARGSERLEU 500
Db      1978 AGCAAGCATTCGATTCGGTCCGTCAGAACACAGACAGCTGACCGCCCAAGGTCCTTTTA 2037
QY      501 SERPROLEUHSIARSESERGLISERVALGILUASPAENYRHSITHRSESPHELEUPHEGLY 520
Db      2038 TCTCCATGCAATCGAAGTGGAGCGTGGAGCAATTAACACACAGCTTCTTTTCGGC 2097
QY      521 LEUSERTHRSESGINGILNHSLEUTHRILYSERVALAGLYLEUGLYLEULYSGLYTRPHIS 540
Db      2098 CTTTCACACAGCCAGACACACCTCAAGAGTCTGCTGGCCTTGAAGGCTGGCAC 2157
QY      541 SERASPILEUVALAPROGLINTHRSETHRPROSERLEUTHRSESERTRPTYRPHALA 560
Db      2158 TCGATATCTTGGCCCCCAGACCTCTACCCCTTCCAGACAGCAGCTGTATTTTGC 2217
QY      561 THRGLUSERSERHISPHETYSERVALASERVALIETRYRGLYGLYSERLASERTYRSE 580
Db      2218 ACAGAGTCTCACTTCTACTCTCTCCTACCATCTACGAGAGACAGATCTATTTGTGGCAGCGG 2277
QY      581 ALATYSESESESEGINLEUPROTHRCSYSGILASPGILVALTYRSEVALARGARGARG 600
Db      2278 GCTACAGTGAGCAGCAGCTGCCACTTGCCGAGACCAAGTCTATTTGTGGCAGCGG 2337
QY      601 GINLYSPROSERASPARVALASPSERARGARGSETRPHISGLUGILUSERPROHEGU 620
Db      2338 CAGAAAGCAAGAGACAGAGCTGACCTCGCGGAGACTGGCANTGAAGAGAGCCCTTTGAA 2397
QY      621 LYSGLNPHELYSARGSERCYSGILMETGLUPHEGLYGLUSERILEMETSESGILUASN 640
Db      2398 AAGCATTTAAACGAGAAAGCTGCCAAATGGAATTTGAGAGAGCATCATGCAAGAAC 2457
QY      641 ARGSEARGGLUGILUEUGLYLYSVALGILYSEGINSESPHESESGILYSEMETGILU 660
Db      2458 AGGTACCGGAAAGACTGGGGAAGTGGGAGTCAGTCTAGCTTTTCGGGCAGATGAA 2517
QY      661 ILEILEGLVALSER 665
Db      2518 ATCATTTGAGTCTCC 2532

```

RESULT 3
ADA53105 standard, cDNA, 2102 BP.
XX
AC ADA53105;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human coding sequence, SEQ ID 673.
XX
XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Noctropic;
KW Gene therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease; gene; ss.
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
PR 24-JAN-2002; 2002US-0350435P.

XX (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y;
XX
DR WPI; 2003-395539/38.
DR P-PSDB; ADA54744.
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 1; SEQ ID NO 673; 205bp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 2102 BP; 542 A; 546 C; 510 G; 504 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	2102
Score:	472.00	Matches:	662
Percent Similarity:	99.10%	Conservative:	0
Best Local Similarity:	99.10%	Mismatches:	3
Query Match:	70.98%	Indels:	6
DB:	7	Gaps:	0

US-10-029-345A-109 (1-665) x ADA53105 (1-2102)

QY 1 MetAlaHisGluMetCilleglYThrGlnIleValThrGluArgPheValAlaLeuGlu 20
Db 56 ATGGCCCATGAGTGAATGTAATCAATTTGTAATGAGAGTTGGTGGCTGCTGGAA 115
QY 21 SerGlyThrGluValLeuLeuIleApsSerArgProPheValGluTyrAenThrSer 40
Db 116 AGTGAACCGGAAAGGTGCTAAATTGATAGCCGGCCATTGTGGAAATACATCATCC 175
QY 41 HisIleLeuGlnAlaIleAenIleAenCysSerIleLeuMetIleArgAlaGluGln 60
Db 176 CACATTTTGGAAACCATTAATATCACTGCTCAAGCTTAAGAAGCAAGGTGCAACG 235
QY 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleHisIleValAapIleAps 80
Db 236 GACAAAGTGTTAATTACAGAGCTCATCCAGCATTCAGCGAAACATTAAGTTGACATTGAT 295
QY 81 CysSerGlnIleValValValIleTyrApsGlnSerSerGlnApsValAAserLeuSerSer 100
Db 296 TGCAGTCAGAAAGTTGATTAATGATCAAAAGCTCCCAAGATGTTGCTCTCTCTTCA 355
QY 101 AspCysPheLeuThrValLeuLeuGlyValLeuGlnIleValSerPheAenSerValHisLeu 120
Db 356 GACTGTTTCTCACTGACTTCTGGGTAACTGAGAGAGACTTCAACTCTGTTCACTG 415
QY 121 LeuAlaGlyValPheAlaGluPheSerArgCysPheProGlyLeuGlyGluIleValSer 140
Db 416 CTTCGAGGTGGGTTTGTGAGTCTCTCTGTTGTTTCCCTGCTCTGTGAAGAAATCC 475
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAenIleGlyProThr 160
Db 476 ACTCTAGTCCCTACCTGATCTTCTCAGCCTTGCTTACCTGTGCCAATTTGGCCACAC 535
QY 161 Arg-IleLeuProAenLeuTyrLeuGlyCysGlnArgApsValLeuAenIleGluLeu 180
Db 536 CT-AATTTCTTCCAAATCTTATCTTGGCTGCGAGCGAGATGCTCAACAAGAGCTGAT 594
QY 180 e-GlnGlnAenGlyIleGlyTyrValLeuAenAlaSerTyr-ThrCysProLysProAps 199
|||||

Db 595 -CGACGAGAAATGGAGATTGGTATGTGTTAATGCCAGCAA-TACCTGTCCAAAGCTGCAC 652
QY 200 PheIleProGluSerHisPheLeuArgValProValAAspSerPheCysGlyValIle 219
Db 653 TTTATCTCCCGAATCTCATTTCTGCGTGTGCTGGTGAAGACAGCTTTGTGGAATAAT 712
QY 220 LeuProTyrLeuApsIleSerValAapPheIleGluValAlaValAAserAenGlyCys 239
Db 713 TTGCGGTGTTGACAAATCAGTATGTTTATTGAGAAAGCAAAAGCTCTCAATGATGT 772
QY 240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAryIle 259
Db 773 GTTCTAGTACATGTTTATGATGATCTCCCGCTCGGCCACCATGCTATCGCTACATC 832
QY 260 MetIleValMetApsMetSerLeuApsGluAlaIleTyrArgPheValIleGlyValApsPro 279
Db 833 ATGAAGAGATGACATGTCTTATGATGAAGCTTACAGATTTGTGAATAAATAAACCT 892
QY 280 ThrIleSerProApsPheAenPheLeuGlyGlnLeuLeuApsTyrGlyValIleVal 299
Db 893 ACTATATCTCCAAATCTTCAATTTTCTGGCCAAATCTTGGACTATGAGAAAGATTAAG 952
QY 300 AenGlnThrGlyAlaSerGlyProLysSerIleValLeuValAAserGluValPro 319
Db 953 AACCAAGACTGAGCATCAGGCGCAAGCAAACTCAAGCTGCTGCACCTGGAGAGCCCA 1012
QY 320 AenGluProValProAlaValSerGlyGlyGlnIleValSerGluThrProLeuSerPro 339
Db 1013 AATGAACCTGTCTCTGCTCTCTCAGAGGTGACAGAAAGAGAGCGCCCTCACTCCA 1072
QY 340 ProCysAlaApsSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSer 359
Db 1073 CCTGTGCGAGCTGTGACTCTCAGAGCGAGAGCAAGGCGCTGATCCCGCAGC 1132
QY 360 ValProSerValProSerValGlnProSerLeuLeuGluApsSerProLeuValGlnAla 379
Db 1133 GTGCCAGCGTCCAGCGGTGACGCGTGTGTTAGAGACAGCCGCTGGTACAGGGG 1192
QY 380 LeuSerGlyLeuHisLeuSerAlaApsArgLeuGluApsSerAenIleValAAser 399
Db 1193 CTCAGTGGGCTCACCCTGTCCAGAGCTGAGAGAGCAAGCAATTAAGCTCAAGCCTTCC 1252
QY 400 PheSerLeuApsIleLeuSerValSerTyrSerAlaSerMetAlaAAserLeuHisGly 419
Db 1253 TTCTCTGATATCAAAATCAGTTTCAATTAAGCCAGATGAGCAATCTTACATGGC 1312
QY 420 PheSerSerSerGluApsAlaLeuGluTyrTyrIlePProSerThrThrLeuApsGlyThr 439
Db 1313 TTCTCTGATATCAAAATCAGTTTCAATTAAGCCAGATGAGCAATCTTACATGGC 1372
QY 440 AenIleValCysGlnPheSerProValGlnIleLeuSerGluGlnThrProGluThrSer 459
Db 1373 AACCAAGCTATGCAAGTTCTCCCTGTTCAAGAACTATCGAGACAGACTCCCGAAACAGT 1432
QY 460 ProApsIleGluGlnAlaAAserIleProIleValLeuGlnThrAlaArgProSerApsSer 479
Db 1433 CCGATTAAGAGAGAGGCGAGATCCCAAGAGCTGACAGCTGAGGCTTCAAGCACG 1492
QY 480 GlnSerIleArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 499
Db 1493 CAGAGACAGCATTTGATTCGATCAGAACACAGACAGAGTGGCACCGCCAGAGTCCCTT 1552
QY 500 LeuSerProLeuHisArgSerGlySerValGluApsAenTyrHisThrSerPheLeuPhe 519
Db 1553 TTATCTCCACTCATCAGAGTGGAGCGTGGAGGCAATTAACACACAGCTTCCCTTTC 1612
QY 520 GlyLeuSerThrSerGlnIleLeuThrIleValSerAlaGlyLeuGlyValIleVal 539
Db 1613 GGCCTTTCCACAGCGACGAGCACTCACAGAGTGTGCTGGGCTTAAAGGCTGG 1672
QY 540 HisSerApsIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrTyrPhe 559
Db 1673 CACTGGATATCTTGGCCCCCGAGACTTACCCCTTCCCTBACAGAGAGCTGTATTTT 1732

QY 560 AlathrgIuSerSerHisphETyrSerAlaSerAlaIleTyrGlySerAlaSerTyr 579
 DB 1733 GCCACAGAGTCTCTACCTTACTCTGCTCCAGCCATCTACCGAGGCACTGCGAGTTAC 1792
 QY 580 SerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 599
 DB 1793 TCTGCTACAGACTGCACCGAGCTGCCACTTGGCGAACAAGCTTATTCTTGCGCGAG 1852
 QY 600 ArgGlnIyProSerArpaAlaaspSerArgArgSerTphIsGluIuSerProPhe 619
 DB 1853 CGGAGAAAGCCAAATGACAGAGCTCCCGCGGAGCTGCGCATGAGAGAGACCCCTTT 1912
 QY 620 GluIySerInPheIySArGArSerCysGlnMetGluPheGlyGluSerIleMetSerGln 639
 DB 1913 GAAGAAGAGTTTAAACCAAGAGCTGCCAAATGAAATTGAGAGAGCATCATGTACAG 1972
 QY 640 AsnArgSerArgGluIuSerGlyIyValGlySerGlnSerSerPheSerGlySerMet 659
 DB 1973 AACAGGTCACGCGAAGAGCTGGGGAAGTGGGAGTCAGTCTTTCGGGCAACATG 2032
 QY 660 GluIleIleGluValSer 665
 DB 2033 GAAATCATTTAGGCTCTCC 2050

RESULT 4
 AAD09492
 ID AAD09492 standard; DNA; 2732 BP.

AC AAD09492;
 DT 10-SEP-2001 (first entry)
 DE Human SGP002 phosphatase polypeptide encoding DNA.

XX Human; SGP002 phosphatase polypeptide; phosphatase-related disease;
 KW immune-related disorder; organ disease; organ transplant rejection;
 KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
 KW metabolic disorder; hematopoietic cancer; mood disorder; cardiact;
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
 KW attention disorder; cognition disorder; psychotic disorder; cystostatic;
 KW neurological disorder; virutide; nootropic; cerebroprotective; therapy;
 KW neuroprotective; antibacterial; vulnerrary; tranquilliser; antiaethmatic;
 KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hyperensive;
 KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
 KW MKP; migraine; chromosome 12p11.1-p12.1; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH 538..2535
 FT CDS /tag= a
 FT /product= "Human SGP002 phosphatase polypeptide"

XX MO200146394-A2.

XX 28-JUN-2001.

XX 21-DEC-2000; 2000MO-US034736.

XX 21-DEC-1999; 99US-0173255P.

XX 28-DEC-1999; 99US-0175766P.

XX 25-JAN-2000; 2000US-0178078P.

XX 31-JAN-2000; 2000US-0179301P.

XX (SUGEN) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Manning G, Sudareanam S;
 PI Hill RJ, Flanagan P;
 XX WPI: 2001-418058/44.
 DR P-PSDB; AAE04834.

XX Novel phosphatase polypeptide useful for treating cancers, immune-related
 PT diseases and disorders, cardiovascular disease, brain or neuronal-
 PT associated diseases and metabolic disorders.

XX Claim 29; Fig 1; 186pp; English.

CC The present invention relates to phosphatase polypeptides, nucleotide
 CC sequences encoding them, as well as various products and methods useful
 CC for the diagnosis and treatment of various phosphatase-related diseases
 CC and conditions. Substance that modulates the activity of phosphatase
 CC polypeptide is used to treat immune-related diseases and disorders,
 CC cardiovascular disease, brain or neuronal-associated diseases and
 CC metabolic disorders, including cancers of tissues, cancers of
 CC haematopoietic origin, diseases of central and peripheral nervous system,
 CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
 CC lateral sclerosis, viral infections, infections caused by prions,
 CC bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual
 CC dysfunction, mood disorders, attention disorders, cognition disorders,
 CC hypotension, hypertension, psychotic disorders, neurological disorders,
 CC dyskinesias and organ transplant rejection. The present sequence is a DNA
 CC encoding human SGP002 phosphatase polypeptide. This sequence is
 CC classified as dual specificity phosphatase (DSP) and MAP kinase
 CC phosphatase (MKP). SGP002 gene maps to chromosomal position 12p11.1-p12.1
 XX

SQ Sequence 2732 BP; 710 A; 684 C; 686 G; 652 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:
472..00	2732	472.00	663
Percent Similarity:		99.40%	0
Best Local Similarity:		99.40%	2
Query Match:		70.98%	4
DB:		Gaps:	0

US-10-029-345A-109 (1-665) x AAD09492 (1-2732)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu 20
 DB 538 ATGGCCAGTGAAGATTTGAAGATTCGAATTTGTAAGAGGTTGGCTCTGCTGAAA 597
 QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluIyAsnThrSer 40
 DB 598 AGTGAAGCGAAGAAAGCTGCTATGATAGCCGGCATTTGGGAATCATACATCTCC 657
 QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIyLeuMetIySarArgLeuGln 60
 DB 658 CACATTTGGAAGCATTAATATCACTGCTCCAGCTTATGAAGCGAAGTTGCAACAG 717
 QY 61 AspIyValLeuIleThrGluLeuIleGlnHisSerAlaIyHisIyValAspIleAsp 80
 DB 718 GACAAAGTGTTAATTAACAGAGCTATCAAGCATTCAGCAACATYAAAGTTGACATTGAT 777
 QY 81 CysSerGlnIyValValIyAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB 778 TGCACTCAGAAAGTGTGATTTAGATCAAGATCCCAAGATGTGCTCTCTCTTCA 837
 QY 101 AspCysPheIeuthrValIleLeuGlyIyLeuGluIySerPheAsnSerValHisLeu 120
 DB 838 GACTGTTTCTCACTGATCTTGGGTAACTGGAGAAAGACTTCAACTCTGTTCACTG 897
 QY 121 LeuAlaGlyIyPheAlaGluPheSerArgCysPheProGlyLeuGlyGluIySerSer 140
 DB 898 CTTCGAGGTGGTGTGCTGATGCTCTCTGTTGTTTCCCTGCTGCTGTAAGGAAATCC 957
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProthr 160
 DB 958 ACTCTAGTCCCTACCTGACATTTCTCAGCCTTGCTTACTGTTCCAAACATGGGCCAAC 1017
 QY 161 ArgIleLeuProAsnLeuIyTrLeuGlyCysGlnArgAspValLeuAsnIyGluLeu 180
 DB 1018 CGAATCTTCCCAATCTTATCTTGCTGCGAAGAGATGTCTCTCAACAAAGAGCTGAT- 1076

QY 181 -GlnGlnaNGlyIleGlyTyrrValleuanaIaSerTyr-ThrCysProIysProAsp 200
 DB 1077 GGAGCAGAAATGGAGTGGTTATGTTAATGCCAGCA-TACTCTCCAAAGCCGACT 1135
 QY 200 heIleProGluSerHisPheLeuArgValProValIenIaPsePheCysGlyIleIle 220
 DB 1136 TTATCCCGAGTCTCATTTCTCGCTGCTGTGAATGACAGCTTTGTGAGAAATTT 1195
 QY 220 euProIleuAspIysSerValIaPheIleGlyIleValIaIaSerAenGlyCysV 240
 DB 1196 TGCCGGTGTGAGCAATCATGATTTCAATGAGAAAGCAAAAGCCCTCCAAATGATGG 1255
 QY 240 alleuValHisCysLeuAlaGlyIleSerArgSerIaThrIleAlaIleAlaTyrIle 260
 DB 1256 TTCTATGCACTGTTTACGTGGGATCTCCGCTCCGCCACCATCGCTATCGCTAATCA 1315
 QY 260 etIlyArgMetAspMetSerLeuAspGluAlaTyrArgPheValIlyGlyIlyArgPro 280
 DB 1316 TGAAGAGATGACATGCTTTAGATGAAGCTTACAGATTGTGAGAAAGAAAGACCTA 1375
 QY 280 hrlIleSerProaenPheAenPheLeuGlyGlnIleuAenPyrGlyIlyIleIlyA 300
 DB 1376 CTATATCTCCAACTTCAATTTCTGAGCCACTCTGAGCTATGAGAAAGAAATTTA 1435
 QY 300 sngIlnThrGlyIaSerGlyProIysSerIysLeuIysIleuHisIleGlyIlyPro 320
 DB 1436 ACCAGACTGGAGCATAGGGCCAAAGCAAACTCAAGCTGCTGCACTGGAGAACCA 1495
 QY 320 sngIubProValProAlaValSerGlyIlyGlyIlyIysSerGluThrProIysSerPro 340
 DB 1496 ATGAACTGCTGCTGCTGCTGCTGAGAGGTGAGCAGAAAGAGAGAGAGCCCTCAGTCCAG 1555
 QY 340 rocYalIaAspSerIaThrSerGluAlaAlaGlyIlyIlyIlyIlyIlyIlyIlyIly 360
 DB 1556 CCTGTGCCACTGTGTACTCTGAGAGCAGAGCAAAAGCCCGTGCATCCCGCAGAG 1615
 QY 360 aIProSerValProSerValGlnProSerIleuGluIaAspSerProIeValIaGlnAl 380
 DB 1616 TGCCAGCTGCCCCAGCCTGTCAGCCCTCGCTTTAAGACAGCCCGCTGCTAGAGCCG 1675
 QY 380 euseGlyIleuHisIleuSerIaAspArgIleuGluAspSerAenIlyIleuIlyArgSer 400
 DB 1676 TCAGTGGCTGACACTGTCGCGAGAGAGGCTGAGAAACAGCAATAGCTCAAGCTTCT 1735
 QY 400 heserIleuAspIleIysSerValSerTyrSerIaSerMetIaAlaSerIleuHisGly 420
 DB 1736 TCCTCTGATATCAATCATGTTTCATATTCAGCCAGCATGGCAGCATCTTACATGGCT 1795
 QY 420 heserSerSerGluAspAlaIleuGluTyrTyrIysProSerThrThrIleuAspGlyTh 440
 DB 1796 TCTCTCATCAGAAATGCTTTGGAAITACTCAAAACCTTCCACTTCTGGATGGAGCA 1855
 QY 440 sulyIleuCyGlnIlePheSerProValGlnIleuSerGluIlnThrProGluThrSer 460
 DB 1856 ACAAGCTATGCCAGTCTCCCTCTGTCAAGAACTATCGAGAGAGACTCCCGAAACCAAGTC 1915
 QY 460 roAspIlyGlyIleuAlaSerIleProIlyIlyIleuGlnThrIaAspProSerAspSer 480
 DB 1916 CTGATTAAGAGAGAGAGCAGCATCCCAAGAAAGCTGAGACTGCGAGCCCTTCAGAGAGCC 1975
 QY 480 IAserIlyArgIleuHisSerValArgThrSerSerSerGlyThrIaGlnArgSerIleu 500
 DB 1976 AAGAGAGAGATTCATTCGGTACAGAACAGCAGCAGTGGACCCCGCAGAGGTCCCTTT 2035
 QY 500 euseProIleuHisArgSerGlySerValGluAspAenTyrHisThrSerPheLeuPhe 520
 DB 2036 TATCTCACTGATCGAAGTGGAGAGCGTGGAGAGCAATTAACACACAGCTTCTTTTGG 2095
 QY 520 IlyIleuSerThrSerGlnIleuHisIleuThrIysSerIaGlyIleuGlyIleuIlyIly 540
 DB 2096 GCTTTTCCACACAGCAGCAGCACTCAAGAAAGTCTGCTGCGCTTAAAGGCTGGC 2155
 QY 540 IAserAspIleuAlaProGlnThrSerThrProSerIleuThrSerSerTyrIlyPheA 560

DB 2156 ACTCGATATCTGGCCCCCGACACTTACCCCTTCCCTGACAGCAGCTGGTATTTG 2215
 QY 560 IatThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrIlyGlySerAlaSerTyr 580
 DB 2216 CCAGAGAGCTTCACACTTCTACTGCTGCTCCAGCATTCACGAGCAGTGCAGTTACT 2275
 QY 580 eAlaTyrSerCysSerGlnIleuProThrCysGlyAspGlnValTyrSerValArgArg 600
 DB 2276 CTGCTTACAGCTGCAAGCCAGCTGCCCATCTTGGAGACCAAGTCTTCTGCGAGGC 2335
 QY 600 rGlnIlySerProSerAspArgAlaAspSerArgArgSerTrpHisGluIleuSerProPhe 620
 DB 2336 GGCAGAGCCAAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2395
 QY 620 IulysGlnPheIlyArgArgSerCysGlnMetGluPheGlyIleuSerIleMetSerGlu 640
 DB 2456 ACAGGTACAGGAGAGAGTCTGGGAGAAAGTGGAGTCAAGTCTTTTGGGAGCATGG 2515
 QY 660 IulIleIleGlyValSer 665
 DB 2516 AATCATTTAGGCTCTCC 2532
 RESULT 5
 AAH9685
 ID AAH9685 standard; cDNA; 2966 BP.
 AC AAH9685;
 DT 16-OCT-2001 (first entry)
 DE Human protein encoding cDNA sequence SEQ ID NO:520.
 KW Human; cancer; HIV infection; human immunodeficiency virus;
 KW antinflammatory; antirheumatic; antidiabetic; immunosuppressive;
 KW antibacterial; endocrine; cardiac; central nervous system; vitamin;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anaemia;
 KW antileptant; haemostatic; vulnery; antidiabetic; osteopathic; eczema;
 KW dermatological; antileptant; antidiabetic; antidiabetic; cytostatic;
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder; ss.
 KW Homo sapiens.
 OS XX
 PN MO200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 22-DEC-2000; 2000WO-US035017.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-457603/49.
 DR P-PSDB; AAM25744.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the

PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.

XX Claim 1; Page 578; 1217pp; English.

XX
CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: antiinflammatory; antineumatic;
CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiac;
CC central nervous system; viricide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnary;
CC antitumor; osteopathic; dermatological; antiallergic; antiallergic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC of disorders associated with the activity of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders

XX Sequence 2966 BP; 809 A; 735 C; 688 G; 733 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 0
Score: 472.00
Percent Similarity: 99.39%
Best Local Similarity: 99.39%
Query Match: 70.98%

Length: 2966
Matches: 648
Conservative: 0
Mismatch: 2
Indels: 4
Gaps: 0

US-10-029-345A-109 (1-665) x AAH99685 (1-2966)

QY 16 ValAlaLeuLeuGluSerGlyThrGluValLeuLeuLeuAspSerArgProPheVal 35
DB 68 GGGGCTCTCTGGAAAGTGGAAAGCGAAAGAGTCTCTAATGATGAGCGGCAATTTGTG 127
QY 36 GluTyrAsnThrSerHisLeuGluAlaLeuAsnGlySerLeuMetLeu 55
DB 128 GAATCAATATACATCCACATTTTGGAAAGCCATTAAATACATGCTCCCAAGCTTATGAAG 187
QY 56 ArgArgLeuGluGlnAspLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 75
DB 188 CGAAGGTTCACACAGCAAGAGTGTATTAACAGAGCTCATCCAGATTCAACGAAACAT 247
QY 76 LysValAspLeuAspCysSerGlnLysValValValValValValValValValValVal 95
DB 248 AAGGTTGACATTGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 307
QY 96 AlaSerLeuSerSerAspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPhe 115
DB 308 GCGCTCTCTCTTCAAGAGCTTTTCTCACTGACTTCTGGTAAACGAGAAAGAGCTTC 367
QY 116 AsnSerValHisLeuLeuValGlyGlyPheAlaGluPheSerArgCysPheProGlyLeu 135
DB 368 AACTCTGTTACCTGCTGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 427
QY 136 CysGluGlyLysSerThrLeuValProThrCysLeuSerGlnProCysLeuProValAla 155
DB 428 TGTGAAGAAATTCACCTCTAGTCTCTACCTGATTTCTCAGCTTCTGCTTACCTGCTCC 487
QY 156 AsnIleGlyProThrArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeu 175
DB 488 AACATTGGGCGCAACCGAATCTCTCCCAATCTTTATCTTGGCTGCACAGAGATGCTCTC 547
QY 176 AsnLysGluLeuIle-GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrC 195
DB 548 AACCAAGAGAGCTGAT-GCAGCAGAAATGGAGTGTATGTATGTAAATGCCAGCAAA-TACCT 605

QY 195 YspProLysProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerP 215
DB 606 GTCCAAAGCCTGACCTTATCCCGAGCTCAATTTCTCTCGTGTGCTGTGAATGACAGCT 665
QY 215 heCysGluLysIleLeuProThrLeuAspLysSerValAspPheIleGluLysAlaLysA 235
DB 666 TTTGGAGAAATTTTGGCGGTGGTGGAGCAATCATGATGATTTTCAATGAGAAAGCAAAAG 725
QY 235 LAserAsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleA 255
DB 726 CTTCAAGATGATGTTCTTACGACACTGTTAGCTGGGATCTCCCGCTCCCGCAACATCG 785
QY 255 LalleAlaTyrIleMetLysArgMetLysMetSerLeuAspGluAlaTyrArgPheVal 275
DB 786 CTATGCTTACATCATGAAAGAGATGACATGCTTTAGATGAAGCTTACAGATTTTGA 845
QY 275 YsgLysArgProThrIleSerProAsnPheAsnPheLeuGlyGluLeuAspTyrG 295
DB 846 AAGAAAAAGACCTATCTATCTCCAACTTCAATTTTCTGGCCAACTCTGAGACTATG 905
QY 295 LulysLysIleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeu 315
DB 906 AAGAAAGATTAAGAACACAGACTGAGACATGAGGCGCAAGAGCAAACTCAAGCTGCTGC 965
QY 315 LsleuGluLysProAsnGluProValProAlaValSerGluGlyGluLysSerGluT 335
DB 966 AACTGGAAGCCAAATGAACCTGTCTCTGTTCAAGAGGTGACAGAAAGCGCAGA 1025
QY 335 hrProLeuSerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProV 355
DB 1026 CGCCCTTACGTCACCGCTGTGCCACTCTGTACTCAGAGCGCAGAGCAAAAGGCCCG 1085
QY 355 AlHisProAlaSerValProSerValProSerValGlnProSerLeuLeuAspSerP 375
DB 1086 TGCAATCCCGCAGGCTGCCAGCTGCGCAGCGTGCAGCGCTGTTAGAGGACAGCC 1145
QY 375 LoleuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerLeu 395
DB 1146 CGCTGGTACAGGCCCTCAGTGAGGCTGCACTGTCCGAGAGAGGCTGGAAGACGCAATA 1205
QY 395 YsLeuLysArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAla 415
DB 1206 AGCTCAAGCGTCTCTCTCTGATATCAATCAGTTTCAATTCAGCCAGATGGCAG 1265
QY 415 LAserLeuHisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrT 435
DB 1266 CATCTTACATGGCTCTCTCTCATCAGAAAGATGCTTGGAAATCATCAAACTTCCACTA 1325
QY 435 hrLeuAspGlyTyrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnT 455
DB 1326 CTCTGGATGGAGCCAAAGCTATGCAAGTTCCTCTCTTCAAGAACTATCGAGAGAGA 1385
QY 455 hrProGluThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAla 475
DB 1386 CTCCGAAACACAGTCCGATTAAGAGSAAAGCCAGCATCCCAAGAGCTCACAACGCCCA 1445
QY 475 rgrProSerAspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrA 495
DB 1446 GGCCTTACAGACAGCAGAGCAAGCATGTCATTCGGTCAGAAACACAGAGAGTGGACCG 1505
QY 495 LAsnArgSerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHis 515
DB 1506 CCGAAGAGTCCCTTTTATCTTCACTGATGAGAGTGGAGCGGAGAGCAATTAACACA 1565
QY 515 hrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuG 535
DB 1566 CCAAGCTTCTTTTCCGGCTTTTCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1625
QY 535 LysLeuLysGlyTyrPheLysSerAlaLeuAlaProGlnThrSerThrProSerLeuThrS 555
DB 1626 GCCTTAAGGAGCTGCACTGGATATCTTGGCCCCCAGACCTCTTACCCCTCTCGAGACA 1685

QY 555 exSerTrpIYrPhealathrGluSerSerHisPheTrYSerAlaSerAlaIleTyGlyG 575
 DB 1686 GCAGCTGATTTTGGCAGAGACTCTCACTTACTTGTGCTCAGCCATCTTAGGAG 1745
 QY 575 lYSerAlaSerTrYSerAlaTySerCySserGlnLeuProThrCySGlyASPGLInValT 595
 DB 1746 GCAGTCCAGTACTCTGCTCCTAAGCAGCCAGCCAGCTCCCACTTGGGAGACCAAGTCT 1805
 QY 595 yfSerValArgaArgGlnIlyeProSerSerPrgAlaAspSerArgaArgSerTrpHisG 615
 DB 1806 ATTCTGGCGCAGCGCAGAACCCAAAGTACAGAGCTGACTCGCGCGAGCTGGCAATG 1865
 QY 615 lUGluSerProPheGlnIlyeGlnPheIlyeArgaArgSerCySGlnMetGluPheGlyGln 635
 DB 1866 AAGAGAGCCCCCTTGAAGACATTTAAAGCAGAACCTCCAAATGGAATTTGGAGAGA 1925
 QY 635 erlMetSerGluAmaArgSerArgGluGluLeuGlyIlyValIlySerGlnSerSerP 655
 DB 1926 GCATCATGTCAAGAACAGGTCAAGGAGAGCTGGGGAAGTGGCGACATCTTAGCT 1985
 QY 655 heSerGlySerMetGluIleIleGluValSer 665
 DB 1986 TTTGGGCGACAGCATGAAATCATTAGGCTCTCC 2017

RESULT 6
 AAS15768
 ID AAS15768 standard; cDNA; 3059 BP.
 AC AAS15768;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE cDNA sequence of human (dual specificity phosphatase) DUSP-10.
 XX
 KW Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke;
 KW neuronal degeneration syndrome; Alzheimer's disease; depression;
 KW schizophrenia; asthma; immune disorder; inflammatory process; arthritis;
 KW osteoporosis; diabetes; gene therapy; chromosome 12; 8s.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 127..2124
 FT /*tag a
 FT /product= "DUSP-10 protein"
 XX
 PN WO200177340-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001MO-EP003966.
 XX
 PR 10-APR-2000; 2000EP-00107143.
 XX
 PA (MERE) MERCK PATENT GMBH.
 XX
 PI Duecker K;
 XX
 DR WPI; 2002-010917/01.
 DR P-PSDB; AAU09946.
 XX
 PT Novel dual specificity phosphatase polypeptides useful for treating
 PT cancer, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's
 PT disease, depression, schizophrenia, asthma and immune disorders.
 XX
 PS Claim 5; Page 34-37; 43pp; English.
 XX
 CC The present invention relates to a new isolated dual specificity
 CC phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid
 CC sequence that is fully defined in the specification. The invention also
 CC provides a sequence encoded by a 3059 nucleotide sequence fully defined
 CC in the specification, and a sequence having at least 95 % identity to the
 CC polypeptide, or fragments or variants of DUSP-10. The invention is useful

CC for treating cancer e.g. leukaemia, colon carcinoma, lung cancer,
 CC prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy,
 CC stroke, neuronal degeneration syndromes, Alzheimer's disease, depression,
 CC schizophrenia, cardiac myopathies, asthma, immune disorders,
 CC inflammatory processes e.g. arthritis, bowel disease, type I diabetes,
 CC osteoporosis, diabetes and diabetes associated diseases. The molecules of
 CC the invention are also useful as vaccines for inducing immunological
 CC response in a mammal, in disease diagnosis and in assays for screening
 CC agonistic or antagonistic compounds. Other uses of the invention include
 CC identifying membrane bound or soluble receptors, as a diagnostic reagent,
 CC in chromosome localization studies, and as a valuable tool in tissue
 CC expression studies. The present sequence represents cDNA of the human
 CC dual specificity phosphatase, DUSP-10. DUSP-10 is located on chromosome
 CC 12
 XX
 SQ Sequence 3059 BP; 831 A; 761 C; 709 G; 758 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 3059
 Score: 472.00 Matches: 663
 Percent Similarity: 99.40% Conservative: 0
 Best Local Similarity: 99.40% Mismatches: 2
 Query Match: 70.98% Indels: 4
 DB: 6 Gaps: 0
 US-10-029-345A-109 (1-665) x AAS15768 (1-3059)
 QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
 DB 127 ATGGCCCATGAGATGATGTAAGTCAATATTGTAAGAGGTTGGTCTGCTGAGAA 186
 QY 21 SerGlyThrGluIlyeValLeuLeuIleAspSerArgProPheValGluIyrAntThrSer 40
 DB 187 AGTGAACGGAAGAAAGTGTCTTAATGATGACCGGCAATTTGGAAATACATATACATC 246
 QY 41 HisIleLeuGlnAlaIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 60
 DB 247 CACATTTTGGAGCCATTATATATCACTCTCCAACTTATGAGGAAAGTTGGCAACG 306
 QY 61 AspIlyeValLeuIleThrGluLeuIleGlnHisSerAlaIlyeValAspIleAsp 80
 DB 307 GACAAAGTGTATTAATACAGAGCTCATCCAGATTCAGCAAAATATAGTTGACATTCAT 366
 QY 81 CySerGlnIlyeValValValIlyeAspIleAsnIleAsnIleAsnIleAsnIleAsnIle 100
 DB 367 TCGAGTCAGAGGTTGATGATTCAGATCAAGCTCCCAAGTGTGCTCTCTCTTCA 426
 QY 101 AspCyPheLeuThrValLeuLeuGlyIlyeValIlyeSerPheAsnSerValHisLeu 120
 DB 427 GACTGTTTCTCACTGTAATCTTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACCTG 486
 QY 121 LeuAlaGlyIlyePheAlaGluPheSerArgCyPheProGlyLeuCySGluGlyIlyeSer 140
 DB 487 CTGGCAGGTGGGTTGTGATGTTCTCTGTTTCCCTGGCCCTGTGAAGGAAATCC 546
 QY 141 ThrIleValProThrCyValIleSerGlnProCySleuProValAlaAsnIleGlyProThr 160
 DB 547 ACTCTAGTCCCTACCTGCAATTCCTCAGCTTCTTCACTGTTGCCAAACATTTGGCCCAAC 606
 QY 161 ArgIleLeuProAsnLeuIlyeGlyCySGlnArgAspValLeuAsnIlyeGluLeu 180
 DB 607 CGAATTTCTCCAAATCTTATCTTGGCTCCAGCAGAGTCTCTCAACAAGAGCTGAT- 665
 QY 181 -GlnGlnAsnGlyIlyeGlyThrValLeuAsnAlaSerIlyeThrCySProIlyeProAsp 200
 DB 666 GCAGCAGAAATGGATGGTATGTATGTTAAATCCAGCAA-TACCTGTCCAAAGCCGACT 724
 QY 200 helIleProGluSerHisPheLeuArgValProValAsnAspSerPheCySGluIlyeIle 220
 DB 725 TTATCCCGCAGCTCATTTCTGCGGTGCTGTGAATGACAGCTTTGTGAGAAATTT 784
 QY 220 euProTrpLeuAspIlyeSerValAspPheIleGluIlyeValIlyeValIlyeValIlyeVal 240

Best Local Similarity: 99.40% Mismatches: 2
 Query Match: 70.98% Indels: 4
 DB: 6 Gaps: 0

US-10-029-345A-109 (1-665) x ABN59704 (1-3104)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
 DB 159 ATGCCCATGATGATGATGAACTCAATGTTACTGAGAGGTGGCTCGCTGGAA 218
 QY 21 SerGlyThrGluIleValIleuLeuIleAspSerArgProPheValGluIleThrSer 40
 DB 219 AGTGAACCGAAAAAGCTGCTTAATTGATAGCCGCGCATTTGGTATCAATACATCC 278
 QY 41 HisIleuGluAlaIleAsnIleAsnCysSerIleuMetIleValArgLeuGln 60
 DB 279 CACATTTTGGAAAGCATTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGGACAG 338
 QY 61 AspIleValIleuIleThrGluLeuIleGlnHisSerAlaIleValIleAspIleAsp 80
 DB 339 GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCAACATAAGTTGACATTTGAT 398
 QY 81 CysSerGlnIleValIleValIleAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
 DB 399 TGAAGTCAGAAAGTTGATGATTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 458
 QY 101 AspCysPheLeuThrValIleuLeuGlyIleValLeuGluIleSerPheAsnSerValHisIleu 120
 DB 459 GACTGTTTCTCATCTGACTTCTGGTAACTGGAGAAAGAGCTTCACTGTTCACTG 518
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlySer 140
 DB 519 CTTCAGAGTGGGTTGCTGAGTTCTCTCGTTGTTTCCCTGCTGCTGTAAGAAATCC 578
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
 DB 579 ACCTGATGCCCTACCTGATCTTCTGAGCTTCTGCTTCACTGCTTCAATGGGCAACC 638
 QY 161 ArgIleLeuProAsnLeuIleValIleGlyCysGlnArgAspValIleuAsnIleGlyIle 180
 DB 639 CGAATTTCTTCCAACTTTATCTTGGCTGACAGAGATGTCTCAACAAAGAGCTGAT- 697
 QY 181 -GlnGlnAsnGlyIleGlyIleValIleuAsnAlaSerTyr-ThrCysProIleProAsp 200
 DB 698 GCAGCAGAAATGGATGGTATGTTAATGTCAGAA-TACTGCTCCAAACCTGACT 756
 QY 200 IleIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlyIleIle 220
 DB 757 TTATCCCGAGCTCAATTTCTGCTGCTGCTGTAATGACAGCTTTTGGAAATATT 816
 QY 220 euProIlePheAspIleSerValIlePheIleGluIleValIleValIleSerAsnGlyCys 240
 DB 817 TGCCTGGTGGCAAAATCGATGATTTCAATTGAAAGCAAAAGCTCCAAATGGATGG 876
 QY 240 AlIleValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 260
 DB 877 TTCTATGACACTTTTACGTGGATCTCCGCTCCGACCATGCTATGCTTACATCA 936
 QY 260 eIleValArgMetSerLeuAspGluAlaTyrArgPheValIleGlyIleValIleVal 280
 DB 937 TGAAGAGGATGACATGCTTTAGATGAGCTTACAGATTGTGAAAGAAAAAGACCTA 996
 QY 280 hIleSerProAsnPheAsnPheLeuGlyIleLeuLeuAspTyrGlyIleValIleVal 300
 DB 997 CTATATCTCCCACTCAATTTTCTGGCCCACTCTGACCTATGAAAGAAATTTAA 1056
 QY 300 snGlnThrGlyIleSerGlyProIleSerIleuIleValIleuHisIleuGluIlePro 320
 DB 1057 ACCAGACTGAGATCAAGGCCCAAGCAAACTCAAGCTGCTGCACTGGAGAACCCA 1116
 QY 320 snGluProValProAlaValSerGlyGlyIleValIleValIleSerGluThrProLeuSerPro 340
 DB 1117 ATGAACCTGCTGCTGCTGATCAGAGGTGACAGAAAGCAAGAGAGCCCTCACTGCTCA 1176

QY 340 roCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerV 360
 DB 1177 CCTGTCCGACCTCTGTTACTACAGAGCAGCAGCAAAAGCCCGTGCATCCGCCAGCC 1236
 QY 360 aIleProSerValProSerValGlnProSerLeuLeuIleAspSerProLeuValGlnAla 380
 DB 1237 TGCCACGCTGCCAGCAGCTGAGCCGCTGTTAGAGACAGCCGCTGTACAGGCCG 1296
 QY 380 euSerGlyLeuHisIleuSerAlaAspArgLeuGluAspSerAsnIleValIleVal 400
 DB 1297 TCAGTGGCTGACCTGCTCCGACAGAGCTGGAGACAGCATATAGCTCAAGCTTCTC 1356
 QY 400 hSerLeuAspIleIleuSerValSerTyrSerAlaSerMetAlaIleSerLeuHisIle 420
 DB 1357 TCTCTGATATCAATATCATGTTTCTATTTACCCAGCATGGCAGCATCTTACATGGCT 1416
 QY 420 hSerSerSerGluAspAlaLeuGluTyrTyrIleAspSerThrThrLeuAspGlyThr 440
 DB 1417 TCTCTCATCAGAAAGATGCTTTGAAATACTACAAACCTTCCACTGATGGAGACCA 1476
 QY 440 snIleLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSer 460
 DB 1477 ACAAGCTATGACAGTCTCCCTGTTACAGAACTATCGAGAGAGACTCCGAAACCACTG 1536
 QY 460 roAspIleGluGluAlaSerIleProIleValLeuGlnThrAlaArgProSerAspSer 480
 DB 1537 CTGATAGAGAGAGAGCAGATCCCAAGAGCTGAGAGCCGCTTACAGACAGCC 1596
 QY 480 InSerIleArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 DB 1597 AGAGCAGAGGATTTGATTTGTCAGAACACAGCAGATGGACCCGCCAGAGTCCCTTT 1656
 QY 500 euSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPhe 520
 DB 1657 TATCTCAGCATGATGAGAGGTGGAGGTGAGAGACATTAACACACACACTTCTTTG 1716
 QY 520 IleuSerThrSerGlnGlnHisIleuThrIleSerAlaGlyIleuGlyIleuValIle 540
 DB 1717 GCTTTTCCACAGCAGCAGCAGCTTCAAGATCTGCTGCTGCTTAAAGGCTGGC 1776
 QY 540 IeSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPhe 560
 DB 1777 ACTCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGACTGATTTTG 1836
 QY 560 IeThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyIleSerAlaSerTyr 580
 DB 1837 CCACAGAGCTTCAACACTTCTACTGCTGCTGACCATTAAGAGGACAGTCACTTACT 1896
 QY 580 eAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
 DB 1897 CTGCTTACAGCTGACAGCCAGCTGCCCATTTGGGAGAACCAATCTATTCTTGGCCGAC 1956
 QY 600 rGlnIleProSerAspArgAlaAspSerArgIleSerTyrHisIleGluIleSerProPhe 620
 DB 1957 GGCAGAGCAGATGACAGCTTCACTGCGCGGAGCTGGAGATGAAGAGCCCTTTG 2016
 QY 620 IuIleGlnPheIleArgArgSerCysIleMetGluPheGlyIleuSerIleMetSerGlu 640
 DB 2017 AAAACAGATTTAAAGCAGAAAGCTGCCAAATGGAATTTGAGAGAGCATCATGTAGAG 2076
 QY 640 snArgSerArgGluIleuGlyIleValIleValIleSerGlnSerSerPheSerGlySerMet 660
 DB 2077 ACAGGTACAGGAGAAAGCTGGGAGAAAGGAGCATGATGATCTTTTGGGACAGATGG 2136
 QY 660 IuIleIleGluValSer 665
 DB 2137 AATATATTGAGTCTCC 2153

RESULT 8
 ABR48378
 ID ABR48378 strand: cdna: 3332 BP.
 XX

QY 574 GLYGLYSERLASERTYRSERLALTYRSERCVSESGLEUPEPROTHRCYSGIYAAPGIN 593
DB 2117 GGAGGCGAGTCCAGTTACTGCTTCAAGCTGACGACCTGCGCACTTGCGGAGACCAA 2176
QY 594 VALTYRSERVALARGARGINLYSPROSERASPARGNALASPERARGATSERTRP 613
DB 2177 GCTATTCTGTGGCGAGCGGGCGAGAACCAAGTACAGAGCTACTCGCGCGGAGCTGG 2236
QY 614 HIEGLIGLUSERPROPHGLULYGLINPHELYARGARGSERCYSGINMETGLUPHEGLY 633
DB 2237 CATGAAGAGAGCCCTTTGAAAAGCAATTAAACGACAGAGCTGCCAAATGGAATTTGGA 2296
QY 634 GLUSERLMEWSESGLEUENARGSERARGGLULGLULGLYVALGLYSESGINSE 653
DB 2297 GAGAGCATCTGTGTCAGAGAACAGCTCCGGAAGAGCTGGGGAAGTGGGCACTCAGTCT 2356
QY 654 SERPHESEGLYSEMERGLULLEIIEGLUVALSER 665
DB 2357 AGCTTTTGGGCGACATGGAATCATTTAGAGTCTCC 2392
RESULT 9
ABK47596
ID ABK47596 standard; cDNA; 3496 BP.
XX
AC ABK47596;
XX
DT 02-UTL-2002 (first entry)
XX
DE cDNA encoding human dual-specificity phosphatase-3 (DSP-16) protein.
XX
KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;
KW cancer; graft-versus-host disease; allergy; metabolic disease;
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;
KW intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 562..2559
FT /tag= a
FT /product= "Human dual-specificity phosphatase-3 (DSP-16)
FT protein"
FT
PN WO200226997-A2.
XX
PD 04-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US030124.
XX
PR 26-SEP-2000; 2000US-0235487P.
XX
PA (CEPR-) CEPTYR INC.
XX
PI Lucie RM, Wei B;
XX
DR WPI; 2002-315802/35.
XX
P-PSDB; AAU79156.
XX
PT New DSP-16 polypeptide, useful for identifying modulators of its
PT activity, which can be used in the treatment of disorders such as
PT Duchenne muscular dystrophy, or cancer.
XX
PS Claim 7; Fig 1; 87pp; English.
XX
CC The present invention relates to a new polypeptide, DSP-16, having a 665
CC amino acid sequence, given in the specification, or a variant having at
CC least 50 % identical residues, which retains the ability to
CC phosphorylate an activated mitogen-activated protein (MAP) kinase. The
CC invention can be used for identifying agents which modulate DSP-16
CC activity, for modulation of a proliferative response in a cell, survival

CC of a cell, or differentiation of a cell. The cell displays contact
CC inhibition of cell growth or anchorage independent growth and may display
CC altered intercellular adhesion. The agent may modulate apoptosis, or the
CC cell cycle. The identified modulators can be used to treat Duchenne
CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune
CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal
CC cell proliferation, and cell cycle abnormalities. The present nucleic
CC acid sequence is that of the human dual-specificity phosphatase-3 (DSP-
CC 16) gene located on chromosome 12p. This sequence encodes the human DSP-
CC 16 protein of the invention
XX
SQ Sequence 3496 BP; 939 A; 854 C; 856 G; 847 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0 Length: 3496
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 6 Gaps: 0
US-10-029-345A-109 (1-665) x ABK47596 (1-3496)
QY 1 MetAlaHieGLuMeRileGLYThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 562 ATGGCCCATGAGATGATTTGGAATCTCAATTTGATTCAGAGGTGGTGGCTCTGTCGAA 621
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 622 AGTGGAAACGGAAGAAAGTGGTCTGCTAAATGATACCCGGCCATTGTGGAAATACATACATCC 681
QY 41 HieIleLeuGluAlaIleAsnIleAsnCysSerIleValMetIleArgArgLeuGlnGln 60
DB 682 CACATTTTGGAGCCATTAATATCAACGCTCCAACTTATAGAGGAGTTGCAACAG 741
QY 61 AspIleValLeuIleThrGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 742 GACAAAGTTAATTATACAGAGCTCATCCAGCTTCAGGCAACATATAGTTGACATTGAT 801
QY 81 CysSerGlnIleValIleValIleValIleValIleValIleValIleValIleValIleVal 100
DB 802 TGCAGTCAGAAAGTTGATGATTAACGATCAAACTCCCAAGATGTTGCTCTCTCTTCA 861
QY 101 AspCysPheLeuThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
DB 862 GACTGTTTCTCAGCTACTCTGCTGATTAACGAGAAAGCTTCAACTCTGTCACTG 921
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGlnGlyIleSer 140
DB 922 CTTCAGAGTGGGTTTGTCTGAGTTCTCTCTGTTTCCCTGGCCTCTGTAGAGAAATCC 981
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 982 ACTCTAGTCCCTACCTGATTTCTCAGCTTCTTACCTTACCTGCAACATTTGGCCAAAC 1041
QY 161 ArgIleLeuProAsnLeuIleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
DB 1042 CGAATTTCTCCAAATCTTATCTTCTGCTGCCAGCAGAGTCTCTCAACAGAGCTGAT- 1100
QY 181 -GlnGlnAsnGlyIleGlyTyrValIleAsnAlaSerTyr-ThrCysProIleProAsp 200
DB 1101 GCAGCAGATGGGATTTGTTATGTTTAAATCCAGCAA-TTCCTGTCCAAAGCCTGACT 1159
QY 200 HeIleProGluSerIlePheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
DB 1160 TTATCCCGAGTCTCATTTCTCGCGGTGCTGTGAATGACAGCTTTGTGAAATTT 1219
QY 220 euProTTrpLeuAspIleSerValaAspPheIleGluIleAlaIleValaIleSerArgGly 240
DB 1220 TCCCGTGTGGACAACATCAGTATTTTCATTTGAAAGCAAAAGCCTTCAATGATGATGG 1279
QY 240 alleuValHieCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleM 260

Db 1280 TTCAGTGCACGTGTTAGTGGGATCTCCGCTCCGACCATCGCTATGCCCTACATCA 1339
 QY 260 etlysaagmetaspmetserleuaspglualatyargphevallysglulyeargprot 280
 Db 1340 TGAAGAGATGACATGTCCTTTAGATGACATTCAGATTGGAAGAAAAAGACCTTA 1399
 QY 280 hrileserproanpheasnphelenglvglinleuenuaspyrglulylyalaalya 300
 Db 1400 CTATATCTCCAAACTTCAATTTTCTGGGCCCACTCCCTGACATGAGAGAAAGATTAA 1459
 QY 300 snqinthrjglalaserglyporyserlyseuysleuenuhsleuglulyesproa 320
 Db 1460 ACCAGACTGGAGATCAGGGSCCAAGAGCAACTCAAGCTGCTGACCTGGAGAAACCA 1519
 QY 320 snqinluprovalproalavalserglvglyglulyssergluthrproleuserprop 340
 Db 1520 ATGAACTGTCTCTGTCTCAGAGGCTGACAGAAAGCGAGAGAGCCCTTCACGTCCAC 1579
 QY 340 rocyaslaaspserralathrserrglualalaglvglnargprovalhisproalaser 360
 Db 1580 CCTGTGCCACCTCTGCTACTCTCAGAGSCAGAGCAAGAGCCCGTGCATCCCGCCAGCG 1639
 QY 360 alproservalproservalginproserleuenglunasperproleuvalglinalal 380
 Db 1640 TGCCCAAGCGTCCAGCGTGCAGCCGTGCTGTAGAGAGCAGCCGCTGTATACAGGGCG 1699
 QY 380 euserglyleuhsleuseralaspargleuglunasperenlylsleulyahyserp 400
 Db 1700 TCAGTGGGTGACACCTGTCCGACAGAGCTGGAAACAGCAATAGCTCAAGCCGTCTCT 1759
 QY 400 heserleuaspilleyservalserlyseralasermeralalasersleuhsiglyp 420
 Db 1760 TCTCTGTGATATCAATCAAGTTTCAATATTCAGCCAGCATGGCAGATCTTACATGCT 1819
 QY 420 heserserserglunaspalaleuglutyrylysproserrthrhleuaspgllythra 440
 Db 1820 TCTCCCATCAGAAAGATGCTTGGAACTACCAAACTCCACTGCTGATGGAGACA 1879
 QY 440 snlyslencysglnpbeserprovalglngluleuserglunthrprogluthrserp 460
 Db 1880 ACAAGCTATGCAAGTCTCCCTGTCAGAGAACTATCGAGAGAGACTCCCGAAACAGAGC 1939
 QY 460 roasplysglunualaserileprolyslsleuglunthralaargprosearaspserg 480
 Db 1940 CTGATAGAGAGAGAGCAGACATCCCAAGAGACTGAGACCCCGAGCCTTACAGACAGCC 1999
 QY 480 lnserylsargleuhsiservalargthrserserserglythralaglnargserleu 500
 Db 2000 AGAGCAAGCATGTCATCGTTCAGAACCCAGACAGTGGCACCGCCAGAGGTCCCTTT 2059
 QY 500 euserproleuhsargserglyservalglunaspasrthrsrphelenupeg 520
 Db 2060 TATCTTCATCTGATCGAAGTGGAGGTGGAGAGCAATTAACACACCAAGCTTCTTTTCG 2119
 QY 520 lyseuserthrserrnglnhisleuthlysseraleglyleuglulyeulysgllytrp 540
 Db 2120 GCCTTTCCACGACCCAGACGACACTCAGAACTGCTGCTGGCCCTTAAGGCTGGC 2179
 QY 540 isersaspilleuualaproglinthrserthproserleuthrsersertpyrpha 560
 Db 2180 ACTCGGAATATCTGGCCCCCAGACCTTACCCCTTCCCTGACGAGCTGGTATTTTGG 2239
 QY 560 latthrgluserserhisphetyrservalaseralellythrglylyserlasertys 580
 Db 2240 CCACAGAGTCTCAACACTTCTACTCTGCTCAGCCATCTACGAGAGCGAGTCCAGTAACT 2299
 QY 580 eeralatysercysserglnleuprothrcysglyaspglinaltyrservalargaga 600
 Db 2300 CTGCCTAACAGCTGACGACGAGCCACTTGGCGAGACCAAGTCAATCTGTGGCCAGCC 2359
 QY 600 rglglnlypserasparalaspasrargargsertrphsiglunserpropheg 620
 Db 2360 GGCAGAAAGCAAGTACAGAGCTGACTCGCGCGAGCTGGCATGAAGAGAGCCCTTTG 2419

QY 620 julyeGlnpheylsargargsercysglnmetglupheglyclunserilewetserglun 640
 Db 2420 AAAACAGATTAAAGCAGAAAGCTGCCAAATGGAATTGGAGAGAGCATCATGTACAGA 2479
 QY 640 snargserargglungluleuglulyysvalglyserglnserpheserserglysermetc 660
 Db 2480 ACAGSTCACGGGAAGCTGGGGAAAGTGGGCAGTCACTTATCTTTTGGGCGACATGG 2539
 QY 660 jullellegluvalser 665
 Db 2540 AAATCATTTAGGCTCTCC 2556
 RESULT 10
 AASI4639
 ID AASI4639 standard, cDNA, 3544 BP.
 XX
 AC AASI4639;
 DT 18-DEC-2001 (first entry)
 XX
 DE Human cDNA encoding dual specificity phosphatase 21117.
 KW Human; ss; dual specificity phosphatase 21117; hepatotropic; cytostatic;
 KW hematopoietic disorder; autoimmune disorder; diabetes mellitus;
 KW rheumatoid arthritis; multiple sclerosis; Crohn's disease;
 KW liver disorder; erythroid associated disorder; haemolytic anaemia;
 KW cellular proliferative; differentiative disorder; leukaemia;
 KW acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 589..2586
 FT /tag= a
 FT /product= "Dual specificity phosphatase 21117"
 XX
 PN WO200173059-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009477.
 XX
 PR 24-MAR-2000; 2000US-0191858P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Meyers RA;
 XX
 DR WPI: 2001-611635/70.
 DR P-PSDB; AAU09016.
 PT New human dual specificity polypeptides and nucleic acids for diagnosis
 of disease and treatment of e.g. liver disorders.
 XX
 PS Claim 1; Fig 1; 143bp; English.
 CC The invention relates to two novel human dual specificity phosphatases
 designated 21117 and 38692, the nucleic acids encoding them (including
 CC fragments, allelic variants, their complements or nucleic acids that
 CC hybridise to them) and antibodies raised against the proteins. The
 CC antibody is useful for detecting the presence of the polypeptide, and the
 CC nucleic acid fragments are useful for detecting the presence of the
 CC nucleic acid, for diagnosing disease. Agents (e.g. the antibody and
 CC antisense sequences) are useful for modulating the activity or expression
 CC of the polypeptide/nucleic acid, useful for the treatment of e.g.
 CC hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,
 CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others
 CC listed in the specification) liver disorders, erythroid associated
 CC disorders (e.g. haemolytic anaemia) cellular proliferative or
 CC differentiative disorders, leukaemias (e.g. acute myeloid leukaemia),
 CC carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and
 CC 38692 are also useful for modulating the proliferation, survival,

CC migration or differentiation of a 38692 or 21117-expressing cell. The
CC polypeptide and nucleic acids are useful for identifying modulating
CC agents. The present sequence encodes the dual specificity phosphatase
CC 21117

XX Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3544
Score:	472.00	Matches:	663
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	70.98%	Indels:	4
DB:	5	Gaps:	0

US-10-029-345A-109 (1-665) x AAS14639 (1-3544)

```
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgPheValAlaLeuLeuGlu 20
DB 589 ATGGCCCATGAGATGATGGAATCAATTGTTACTGAGAGAGTGGTGGCTCTGCTGGAA 648
QY 21 SerGlyThrGluIleValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
DB 649 AGTGAACCGAAGAAAGTGGCTGATTTGATAGCCGGCCCAATTGTGTGAATACAAATACATCC 708
QY 41 HisIleLeuGluAlaIleAsnIleAsnCySerIleLeuMetIleArgArgLeuGlnGln 60
DB 709 CACATTTTGAGAGCCCATTAATATCACTGCTCCAGCTTAATGAAGGAAAGTTGCAACAG 768
QY 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
DB 769 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGCGAACAATAGGTTGACATTGAT 828
QY 81 CySerGlnIleValIleValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 829 TGCAGTCAGAGAGGTGTAGTTACGATCAAGATCCCAAGATGTGCTCTCTCTTCA 888
QY 101 AspCyPheLeuThrValLeuLeuGluIleValLeuGluIleValLeuSerPheAsnSerValHisIleu 120
DB 889 GACTGTTTCTCATCTACTCTGAGGTAAACGTGGAGAGAGCTTCAACTCTGTTCACCTG 948
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCyPheProGlyLeuCyGluGlyIleValSer 140
DB 949 CTTCGAGGTGGTGTGCTGAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATATCC 1008
QY 141 ThrLeuValProThrCySileSerGlnProCySleuProValAlaAsnIleGlyProThr 160
DB 1009 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTCACTGTTGCCAATGGGCCAACCC 1068
QY 161 ArgIleLeuProAsnLeuTyrIleGluIleCyGlnArgAspValLeuAsnIleGluIle 180
DB 1069 CGAATTCCTCCCAATCTTATCTGTGGCTGCACGCGAGATGCTCTCAACAGAGGCTGAG- 1127
QY 181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCySProlAspProAsp 200
DB 1128 GCAAGCAATGGAGTTGGTTATGTATATGTTAAATGCCACAA-TACCTGTCCAAAGCCTGACT 1186
QY 200 HisIleProGluSerHisPheLeuArgValProValAsnAspSerPheCyGluIleIle 220
DB 1187 TTATCCCGAGTCTCATTTCTCGTGTGCTGTGATGACAGCTTTGTGAGAAATTTT 1246
QY 220 euProTrrLeuAspIleValSerValAspPheIleGluIleValAlaValAsnAsnGlyCyav 240
DB 1247 TGGCGGTGGTTGGCAATCATGATGATTTCAATTAGAAAGCAAAAGCCTCAATGAGATGG 1306
QY 240 alLeuValHisCySleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 260
DB 1307 TTCTTAGTCACTTTTAACTGAGGATCTCCGCTCCGACCATGCTTATGCTTAAATCA 1366
QY 260 euIleArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluIleValArgPro 280
DB 1367 TGAAGAGATGACATGTCTTTAGATGAGCTTACAGATTTGTGAAAGAAAGAAAGACCTTA 1426
```

```
QY 280 HisIleSerProAsnPheAsnPheLeuGluIleLeuLeuAspTyrGluIleValIleVal 300
DB 1427 CTAATCTCCAAACTTCAATTTTCTGGGCCAACTCTGGACATATGAAAGAAATTAAGA 1486
QY 300 enGlnThrGlyAlaSerGlyProIleValSerIleValLeuIleValLeuGluIleValPro 320
DB 1487 ACCAGACTGAGACATCAGGCGCAAGAGCAAACTCAAGCTGTGACCTGGAGAAAGCCAA 1546
QY 320 enGluProValProAlaValSerGluIleGlyIleGluIleValSerIleValProIle 340
DB 1547 ATGAACCTGTCTCTCTGTCTGAGAGGGTGGACAAAAAGCCAGAGCCCTCAGTCCAC 1606
QY 340 roCyAlaAspSerAlaThrSerGluAlaAlaGlyIleArgProValHisProAlaSerV 360
DB 1607 CTTGTGCCACTCTGCTACTACCTAGAGGCGAGCGCAAGAGCCCTGCTATCCCGCAGCG 1666
QY 360 alProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAla 380
DB 1667 TGCACAGGCGTCCAGCGCGTGCAGCCGTGCTTTAGAGACAGCCGCTGTACAGGCGC 1726
QY 380 euSerGlyLeuHisIleSerAlaAspArgLeuGluAspSerAsnIleValIleValArg 400
DB 1727 TCAGTGGGCTGCACCTGTCCGACAGAGCTGGAAGACGCAATAGCTCAAGCTTCTCT 1786
QY 400 HisSerLeuAspIleIleValSerValSerTyrSerAlaSerMetAlaIleSerLeuHisGlyP 420
DB 1787 TCTCTGTGATATCAATATCGTTTCAATTTACCGACGATGCGAGCATCTTACATGGCT 1846
QY 420 HisSerSerSerGluAspAlaLeuGluIleTyrTyrIlePProSerThrThrLeuAspGlyThr 440
DB 1847 TCTCTCATCAAGAGATGCTTTGAAATACTACAAACCTTCCACTGCTGTGATGGAGCA 1906
QY 440 enIleValCyGlnPheSerProValGlnIleLeuSerGluIleThrProGluThrSerP 460
DB 1907 ACNAGCTATGCGAGTCTCCCTGTTCAAGAACTTTCGAGAGACATCCCGAAACCAAGTC 1966
QY 460 roAspIleGluIleValSerIleProIleValLeuGlnThrAlaArgProAspSerG 480
DB 1967 CTGATTAAGAGAGAGCCAGCATCCCGAGAGCTGCAGACCCCGAGCTTTACAGACGCC 2026
QY 480 InsSerIleArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
DB 2027 AAGCAGAGGATGTCATTCGTCAGAACAGAGCAGAGCTGCACCCGCGCAGAGTCCCTTT 2086
QY 500 euSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheG 520
DB 2087 TATCTCCACTGCATGGAAGTGGAGCTGGAGAGCAATTAACCAACAGCTTCTTTTGG 2146
QY 520 IleLeuSerThrSerGlnGlnHisIleuThrIleValSerAlaGlyIleuGlyIleuValGlyTrrH 540
DB 2147 GCTTTTCCACGCGCAGAGCAGCATTCACGAACTGTGCTGGGCTTTAAGGCTGGC 2206
QY 540 IAspAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrrPyrPheA 560
DB 2207 ACTCGGATATCTGGCCCCCGGACCTTCAACCTTCCCTGACAGAGCTGATTTTGG 2266
QY 560 IatThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyIleSerAlaSerTrrS 580
DB 2267 CCACAGAGCTCCACACTTCTACTCTGCTCAGCATCAAGAGGAGTGCACAGTTACT 2326
QY 580 euAlaTyrSerCySerGlnLeuProThrCyGlyIleAspGlnAlaTyrSerValArgArg 600
DB 2327 CTGCTTACAGCTGCAGCGCAGCTGCCACTTGGGAGAACCAATCTTATTTCTGCGCAGGC 2386
QY 600 rGlnIlePProSerAspArgAlaAspSerArgSerTrrHisGluGluIleProPheG 620
DB 2387 GGCAGAGCGAAGTACAGAGCTGACTCGCGCGGAGCTGCGCATGAAGAGAGCCCTTTG 2446
QY 620 IuIleGlnPheIleValArgArgSerCyGlnMetGluPheGlyIleGluSerIleMetSerGlu 640
DB 2447 AAAACAGATTTAAAGCAGAGAGCTCCCAATGAAATTTGAGAGAGCATCATGTGAGAGA 2506
QY 640 enArgSerArgGluIleuGluIleValIleGlySerGlnSerSerPheSerGlySerMetG 660
```


QY	360	alProSerValiProSerValgiNpProSerLeuLeuGiuaApSerProLeuValgiNlAl	380
Db	1667	TGCCAGGCGTCCAGCGGTGACCGCTGCTGTTAGAGAGACGCCGCTGTATACAGCGC	1726
QY	380	eusSerGiLeuHiIseuSerAlaApAglLeuGiuaApSerAsnLyLeuLyAArgSerP	400
Db	1727	TCAGTGGGCTGCACCTGTCCCGAGACAGCGCTGGAAAGACAGAAATAGCTCAAGGTTCT	1786
QY	400	heserLeuAspLleLySerValserTySerAlaSerMetAlaIAserLeuHiIsgLYP	420
Db	1787	TCTCTCGATATCAAAATCAGTTTCATATTCAGCCAGCATGGCAGCATCTTATCATGGCT	1846
QY	420	heserSerSerGluAspAlaLeuGiuyTyTyLyProSerThThriLeuAspLyIhA	440
Db	1847	TCTCTCATCAGAAATACCTTTGGAAATCTACAAACCTTCCATCTCTGGATGGACCA	1906
QY	440	snLyLeuCyGlnInPheSerProValGlnIuLeuSerGluGlnThProGluThSerP	460
Db	1907	ACNAGCTATGCCAGTCTTCCTCTGTTCAGAACTATCGAGACAGACTCCGAAACCACTC	1966
QY	460	roAspLyGlnIuAlaSerLleProLyLyLeuGlnThThAlaArgProSerAspSerG	480
Db	1967	CTGATTAAGAGAGAACCCAGCATCCCAAGAAAGCTGACAGCCGACGGCTTCACAGACCC	2026
QY	480	InserLyArgLeuHiIAserValArgThSerSerSerGlyThThAlaGlnAArgSerLeu	500
Db	2027	AGAGCAACGATTCGATTCGGTGCAGAACACACAGACAGAGGACCGCCACAGAGTCCCTTT	2086
QY	500	eusSerProLeuHiIaArgSerGlySerValGluAspAsnTyHsiThSerPheLeuPheG	520
Db	2087	TATCTCCACTGCATGAAAGTGGAGCGTGGAGAGACAATTACCACACAGCTTCTTTTCG	2146
QY	520	lyLeuSerThSerGlnIhiIseuThrLySerAlaGlyLeuGlyLeuLyGlyTyPh	540
Db	2147	GCCTTTCACACAGCCAGCAGACACCTCAAGAACTCTGCGCTGGCCTTAAGGGCTGGC	2206
QY	540	IAserAspLleLeuAlaProGlnThSerThrProSerLeuThSerSerTrpTyPhea	560
Db	2207	ACTCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTACACAGCAGCTGTATTTTG	2266
QY	560	laThGluSerSerHisPheTySerAlaIAserAlaIleTyGlnGlySerAlaSerTyrs	580
Db	2267	CCACAGAGCTCTCACACTTCTACTCTGCTCAGCCATCTACGAGAGCGAGTCCAGTTACT	2326
QY	580	erAlaTySerCySerGlnLeuProThrCyGlyAspGlnValTySerValArgArgA	600
Db	2327	CTGCTTACAGCTGCAGCAGCAGCTGCCCATTTCCGAGACCAAGTCTATTTCTGCGCAGCG	2386
QY	600	rgGlnLyProSerAspAlaAlaApSerArArgSerTerThIsgLyuIserProPheG	620
Db	2387	GCGCAAAACCCAAAGTACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCCTTTG	2446
QY	620	IuLyGlnPheLyArgArgArgSerCyGlnMetGluPheGlyGluSerLleMetSerGluA	640
Db	2447	AAACACAGTTTAAACGACAGAACTGCCCCAAATGGAAATTTGGAGAGACATCTATGTCACAGA	2506
QY	640	snArgSerArgGluGluLeuGlyLyValGlySerGlnSerSerPheSerGlySerMetG	660
Db	2507	ACAGTCAACGGGAAGAGCTGGGGAAAGTGGCGACGTCTAGCTTTTGGGCGACGATGG	2566
QY	660	IuIleIleGluValSer 665	
Db	2567	AAATCATTTGAGGCTCTCC 2583	
RESULT 12			
ABK14474			
ID	ABK14474 standard; cDNA; 3766 BP.		
XX			
AC	ABK14474;		
XX			
DT	08-MAY-2002 (first entry)		
XX	Human protein phosphatase 7 (PP7) cDNA sequence.		

XX	Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
KM	acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
KW	Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
KM	dementia; Parkinson's disease; developmental disorder; Down's syndrome;
KW	cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma
KM	melanoma; myeloma sarcoma; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	538..2535
CDS	/**tag= a
FT	/product= "Protein_phosphatase_7_(PP7)"
XX	
PN	WO200210363-A2.
XX	
PD	07-FEB-2002.
XX	
PF	26-JUL-2001; 2001WO-US023716.
XX	
PR	28-JUL-2000; 2000US-0221679P.
XX	
PR	03-AUG-2000; 2000US-0223272P.
XX	
PR	10-AUG-2000; 2000US-0224309P.
XX	
PR	18-AUG-2000; 2000US-0226728P.
XX	
PR	30-AUG-2000; 2000US-0229254P.
XX	
PR	08-SEP-2000; 2000US-0231366P.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
XX	Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE,
P1	Stewart EA, Gandhi AR, Patterson C, Lee EA, Hatalla AA, Lu DM;
P1	Tribouley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;
P1	Malia NK, Kearney L;
XX	
DR	PMI; 2002-188735/24.
DR	P-PDB; AAU5789.
XX	
PT	New protein phosphatases, useful for diagnosing, treating or preventing
PT	immune system disorders (e.g. Crohn's disease), neurological disorders
PT	(e.g. Parkinson's disease), or cell proliferative disorders (e.g.
PT	cancers).
XX	
PS	Claim 5; Page 114-115; 117pp; English.
XX	
CC	The present invention relates to a new polypeptide, a naturally occurring
CC	amino acid sequence at least 95 % identical to it, a biologically active
CC	fragment of it or an immunogenic fragment of it. The polypeptides,
CC	polynucleotides, agonists and antagonists are useful for diagnosing,
CC	treating or preventing disorders associated with aberrant expression of
CC	protein phosphatases (PP), particularly immune system disorders e.g.
CC	acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
CC	asthma or Crohn's disease, neurological disorders e.g. epilepsy,
CC	Huntington's disease, dementia or Parkinson's disease, developmental
CC	disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
CC	cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
CC	or sarcoma. The present nucleic acid sequence encodes human protein
CC	phosphatase 7 (PP7) which is one of several human protein phosphatases
CC	(AAU75783-AAU75792) of the invention
XX	
SQ	Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;
XX	
Alignment Scores:	
Pred. No.:	0 Length: 3766
Score:	472.00 Matches: 663
Percent Similarity:	99.40% Conservative: 0
Best Local Similarity:	99.40% Mismatches: 2
Query Match:	70.98% Indels: 4
DB:	Gaps: 0

US-10-029-345A-109 (1-665) x ABK14474 (1-3766)

1 MetAlaHisGluMetCilegLyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20

Db	538	ATGCCCATGAGATGATTGGAACTCAATATTGTAAGAGAGTTGGTGGCTGCTGGAA	597
Qy	21	SeGClYthrgCluYsValleuLeuLiLeaSPSerArpProPheValGluYrYanThrSer	40
Db	598	AGTGAACCGGAAAAAGCTGCTAAATTGATAGCCGGCATTTGTGGAAATCAATACATCC	657
Qy	41	HisIleuGluAaIleAnIleAenCySerIysLeuMetLysArgAgluEngIngn	60
Db	658	CACATTTTGGAGCAATTAATATCAACTGCTCCAGCTTATGAAAGGAGGTGGCAACG	717
Qy	61	AspLysValleuIleThrgIuLeuIleGlnHisSerAlaLysHisLysValAspIleAsp	80
Db	718	GACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCCAAACATTAAGTTGACATTGAT	777
Qy	81	CysSerGluYsValValAlaYrYrAspGlnSerSerGlnAspValAspIleAspSer	100
Db	778	TGCAAGTCAGAAAGTGTAGTTAGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA	837
Qy	101	AspCySPheLeuThrValleuLeuGlyLysLeuGluYsSerPheAsnSerValHisLeu	120
Db	838	GACTGTTTCTCACTGACTTCTGGGTAACTGGAAAGAGCTTCAACTCTGTTACCTCG	897
Qy	121	LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCySGluGlyLysSer	140
Db	898	CTTGCAAGTGGGTTGCTGAGTTCTCGTTGTTTCCCTGGCCTGTGAGGAAATCC	957
Qy	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	958	ACTGATGCTCCTAACGCAATTTCTCAGCCTTGCTTACGTTGCCAAATTTGGGCCAAC	1017
Qy	161	ArgIleLeuProAsnLeuThrLeuGlyCysGlnArgAspValleuAsnLysGluLeu	180
Db	1018	CGAATTTCTCCCAATCTTTATCTTGCTGCCAGCGAGATGCTCTCAACAGAGCTGAT-	1076
Qy	181	-GlnGlnAsnGlyIleGlyTyrValleuAsnAlaSerTyr-ThrCysProLysProAsp	200
Db	1077	GCAGCGAATGGGATTTGTTATGTAAATGCCAGCAA-TACGTGTCMAAGCCTGACT	1135
Qy	200	heIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluYsIle	220
Db	1136	TATATCCCAAGTCTCAATTCCTGCGGTGCTGTGATGACAGCTTTGTGAGAAATTT	1195
Qy	220	eupProIleuAspLysSerValAspPheIleGluYsAlaLysAlaSerAsnGlyCysV	240
Db	1196	TGCCGATGTTGGACAAATCAGTAAGATTTCATTGAGAAACAAAGCCTCCATGATGTG	1255
Qy	240	AlleuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleW	260
Db	1256	TTCATGTGACCTGTTAGCTGGGATCTCCGCTCCGCCACCAATCGCTATCGCTACATCA	1315
Qy	260	eLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValYsGluYsAspProT	280
Db	1316	TGAAGAGATGACATGCTTTTAAGTAAGACTTACAGATTGTGAAAGAAAAAACCTTA	1375
Qy	280	hIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluYsLysIleLysA	300
Db	1376	CTATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGGACTATGAGAAAGATTAGA	1435
Qy	300	enGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLysIleuGluYsProA	320
Db	1436	ACCAACTGAGCACTCAGGCGCAAAAGACAACTCAACTGTGCACTGAGAAAGCCAA	1495
Qy	320	enGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProP	340
Db	1496	ATGAACCTGCTCTGCTGCTCAGAGGTGACAGAAAAAGCAGAGCGCCCTCAGTCCAC	1555
Qy	340	roCyAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerV	360
Db	1556	CCTGTGCCGACTCTGCTACCTCAGAGGACAGAGCAAAAGCCCTGTGATCCCGCAGCG	1615
Qy	360	AlProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaT	380

Db	1616	TGCCCAAGCTGCCCAAGCTGACAGCCCTGCTTTAGAGACAGCCCGCTGATACAGGCC	1675
Qy	380	eusSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuYsArgSerP	400
Db	1676	TCAGTGGGCTGACCTGCTCCGACAGACAGCTGGAAGACAGCAATTAAGCTCAAGCTTCT	1735
Qy	400	heserLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyP	420
Db	1736	TCTCTGATATCAAAATCAGTTTCATATTTACGCGACATGGAGCATCTTACATGGCT	1795
Qy	420	heserSerSerGluAspAlaLeuGluYrYrLysProSerThrThreLeuAspGlyThra	440
Db	1796	TCTCTTATCAGAAATCTCTTGGAAATCTAATAAACCCTTCACTACTCTGATGGACCA	1855
Qy	440	snLysLeuCySGlnPheSerProValGlnLysSerGluGlnThrProGluThrSerP	460
Db	1856	ACAAGCTATGCGAGTTCTCCCTGTCAGAACTATCGACAGACAGCTCCGAAACAGATC	1915
Qy	460	roAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerG	480
Db	1916	CTGATTAAGAGAGAACCGACATCCCAAGAGCTGACAGACCGCCAGGCTTTCAGACAGCC	1975
Qy	480	InserysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuL	500
Db	1976	AGAGCAACGATTTGATTCGCTCAGAACACAGACAGACAGCCGCCAGAGTCCCTTT	2035
Qy	500	eusSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheG	520
Db	2036	TATCTCCACTGCATCGAAGTGGAGCGTGGAGGACAAATTACACACAGACTTCTTTTCG	2095
Qy	520	lyLeuSerThrSerGluGlnHisLeuThrLysSerAlaGlyLeuGlyLysLeuYsIleP	540
Db	2096	GCTTTTCCACGACGACGAGCACTTCAAGAGTGTGCTGGCTGGGCTTTAAAGGCTGGC	2155
Qy	540	IsSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPzPheA	560
Db	2156	ACTCGAATATCTTGCCCCCCAGACCTTACCCCTTCGACAGCAGAGCTGATATTTTG	2215
Qy	560	lathrgLysSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTys	580
Db	2216	CCACAGAGTCTTCACTCTTACTCTGCTCAGACCATCTACGAGGACGAGCCAGTTACT	2275
Qy	580	erAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgA	600
Db	2276	CTGCTTCAAGCTGACGACCACTGCCACTTGCGGAGACCAAGTCTATTTGTGTGCGAGGC	2335
Qy	600	rgGlnYsProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSerProPheG	620
Db	2336	GGCAGAAAGCCAAAGTACAGAGCTGACTCGCGGGAGACTGGCATGAAGAGCCCTTTG	2395
Qy	620	lulysglnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMerSerGluA	640
Db	2396	AAAAGAGTTTAAACGAGAAAGCTGCCAAATGAAATTTGAGAGAGCATCATGTACAGAG	2455
Qy	640	snArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetG	660
Db	2456	ACAGGTACCGGAAAGCTGGGAAAGTGGGCACTCAGTCTTACCTTTTGGGCGAGATGG	2515
Qy	660	lulleIleGluValSer	665
Db	2516	AAATCATTTAGAGTCTCC	2532

RESULT 13
ABN83966
ID ABN83966 standard; DNA; 4790 BP.
AC ABN83966;
XX
DT 06-SEP-2002 (first entry)
XX Human gene sequence #13.
DE Human; brain; consil; hippocampus; foetal brain; diagnosis; gene; de.
XX
KM

XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 184..2181
 /*tag= a
XX
XX MO200252005-A1.
XX
XX PD 04-JUL-2002.
XX
XX PF 20-DEC-2001; 2001MO-JP011217.
XX
XX PR 22-DEC-2000; 2000JP-00389742.
XX
XX PA (KAZU-) KAZUSA DNA RES INST FOUND.
XX PA (CELE-) CELESTAR LEXICO-SCI LTD.
XX
XX PI Ohara O, Nagase T, Nakajima D;
XX
XX DR WPI; 2002-500762/53.
XX DR P-PSDB; ABB97946.
XX
XX PT Genes and their expression products cloned from human cDNA libraries for
XX PT treatment and diagnosis of diseases associated with their expression.
XX
XX PS Claim 1(a); Page 111-117; 238pp; Japanese.
XX
XX CC The invention relates to DNA encoding polypeptides directly cloned from
XX CC cDNA libraries originating in adult whole brain, human tonsil, human
XX CC adult hippocampus and human foetal whole brain. Polypeptides and
XX CC polynucleotides of the invention may be used in the investigation of
XX CC differential expression of the DNA sequences in normal subjects and
XX CC disease patients. They may also be used in the production of antibodies,
XX CC oligonucleotide probes and DNA chips for diagnosis and identification of
XX CC drugs for treatment of diseases with which the DNA sequences are
XX CC associated. The sequences given in records ABB93954-ABB93984 represent
XX CC human gene sequences of the invention
XX
XX SQ Sequence 4790 BP; 1293 A; 1064 C; 1152 G; 1281 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
Pred. No.: 0 Length: 4790
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 6 Gaps: 0
US-10-029-345a-109 (1-665) x ABB93966 (1-4790)
QY 1 MetAlHhIeGluMeIleGlyThrgInIleValThrgInArgLeuValAlaLeuGlu 20
DB 184 ATGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 243
QY 21 SerGlyThrgInIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 244 AGTGGAAACGGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 303
QY 41 HisIleLeuGluIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 60
DB 304 CACATTTTGGAAACCATTAATATCACTGCTCCAGCTTATGAAAGCAAGTTCAC 363
QY 61 AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 364 GACAAAGTGTATTAATACAGAGCTCATCCACATTCAGCGAAACATAAGTTGAT 423
QY 81 CysSerGlnIleValValValValValValValValValValValValVal 100
DB 424 TGCAGTCAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 483
QY 101 AspCysPheLeuThrValLeuLeuGlyIleLeuGluIleLeuSerPheAlaHisIleu 120

DB 484 GACTGTTTTCTACCTGACTGCTGAGTAACCTGAGAAAGCAAGCTTCACTCTGTCACCTG 543
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyIleuGlyIleuSer 140
DB 544 CTGGCAGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 603
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAlaIleGlyProThr 160
DB 604 ACTGATGCTCCCTACCTGACATTTCTGACCTTGTCTTCTGCTGCTGCTGCTGCTGCT 663
QY 161 ArgIleLeuProAlaIleuLeuGlyIleuGlyIleuGlyIleuGlyIleuGlyIleu 180
DB 664 CGAATTCCTCCCAATCTTATCTTGGCTGCCAGCAGATGTCTCAACAAGAGCTGAT- 722
QY 181 -GlnGlnGlnGlyIleGlyIleValLeuAlaIleSerIleThrCysProIleProAsp 200
DB 723 GGAGCAGAAATGGATGGTGTATGTATGTATGTATGTATGTATGTATGTATGTATGTAT 781
QY 200 HeIleProGluSerHisPheLeuArgValProValAlaAlaIleSerPheCysGluIle 220
DB 782 TTAATCCCGAGTCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841
QY 220 euProIleuAlaIleuSerValIlePheIleGluIleValAlaIleSerValIleu 240
DB 842 TGGCGTGTGGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901
QY 240 AlLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIle 260
DB 902 TTCTAGTGACATGTTAGTGGGATCTCCCGCTCCCGCAATCGCTATCGCTATCATCA 961
QY 260 eIleValArgIlePheSerIleuAlaIleValIleValIleValIleValIleValIle 280
DB 962 TGAAGAGATGAGCAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1021
QY 280 hrIleSerProAlaPheAlaIlePheLeuGlyIleuLeuAlaIleValIleValIle 300
DB 1022 CTATATCTCCAAATCTCAATTTTCTGGGCCAATCTGTGACATGAGAAAGATTAAGA 1081
QY 300 sngIleThrIleAlaIleSerGlyProIleSerIleuLeuIleuIleuIleuIleu 320
DB 1082 ACCAGACTGAGCAGATCAGGCGCAAGAGCAAACTCAAGCTGCGACCTGAGAGACCCA 1141
QY 320 sngIleuProValProAlaIleValIleSerGluIleGlyIleuSerGluIleuPro 340
DB 1142 ATGAACCTGTCTCTGTGTCTGATGAGAGGTGACAGAAAGCAAGAGCCCTCAAGTCCAC 1201
QY 340 roCysAlaIlePheAlaIleThrSerGluAlaIleGlyIleuIleuProValIleIle 360
DB 1202 CCTGTCCGACCTCTGCTACCTCAGAGCAGCAGAGCAAGGCCCGTGCATCCGCCAGCG 1261
QY 360 alProSerValProSerValGlnProSerLeuLeuIleuIleuIleuIleuIleu 380
DB 1262 TGGCCGAGCTGCCAGCGTGCAGCGCTGCTGTAGAGCAGACCGCTGTACAGCGCC 1321
QY 380 euSerGlyIleuHisLeuSerAlaIlePheIleGluIleuSerIleuIleuIleuIleu 400
DB 1322 TCAGTGGGTGACCTGTCTGCGAGCAGAGCTGGAAGCAGCAATAGCTCAAGCGTTCCT 1381
QY 400 HeSerLeuAlaIleuSerValIleSerIleValIleSerAlaIleSerAlaIleSer 420
DB 1382 TCTCTGTGATATCAAAATGATTTCAATATTCAGCCAGCAGATGCGAGATCTTACATGCT 1441
QY 420 HeSerSerSerGluIleuAlaIleuGluIleuIleuIleuIleuIleuIleuIleu 440
DB 1442 TCTCTCATCAGAAAGTGTGTTGAAATACCAAACTTCCACTGATGATGAGGACCA 1501
QY 440 sngIleuIleuGlnIlePheSerProValGlnIleuSerGluIleuIleuIleuIleu 460
DB 1502 ACAAGCTATGACAGTCTCTCTGTCAGAACTATCGGAGAGAGATCCCGAAACCAAGTC 1561
QY 460 roAspIleGluIleuAlaIleSerIleProIleValIleGlnIleuIleuIleuIleu 480
DB 1562 CTGATAGAGAGAGAGAGAGATCCCGAGAGAGCTGAGAGCTCCGAGAGCTTACAGACGCC 1621

QY 480 InsSerIysArGLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuL 500
DB 1622 AAGAGCAAGCGATTGCTGCTGAGAACAGCAGCGAGTGGCAACCCGCCAGAGTCCCTTT 1681
QY 500 euSerProLeuHisArgSerGlySerValGluAspAsnIyHsrThrSerPheLeuPheG 520
DB 1682 TATCTCCACTGCTCAAGTGGAGCGTGAGAGCAATTACACACACACTTCCTTTTG 1741
QY 520 IYLeuSerThrSerGlnGlnHisLeuThrIysSerAlaGlyLeuGlyLeuGlyTyrPH 540
DB 1742 GCCTTCCACCAAGCCAGCAGCACTCAGCAAGTCTGCGCTGGCCTTTAAGGCGTGCG 1801
QY 540 iSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheA 560
DB 1802 ACTCGGATATCTTGGCCCCCAGACCTTAACCCCTTCCTGACACAGCACTGGTATTTTG 1861
QY 560 lATHrGluSerSerHisPheIYrSerAlaSerAlaIleTyrGlyIysSerAlaSerTyrS 580
DB 1862 CCACAGATCTTCACACTTCTACTCTGCTCAGCCACTTACGAGAGGCACTGCCAGTTACT 1921
QY 580 exAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgA 600
DB 1922 CTGGCTACAGCTGCAGCAGCTGCCACTTGGCGAGACCAAGTCTATCTGTGCGCAGGC 1981
QY 600 rGGLnIysProSerAspArgAlaAspSerArgArgSerTyrPHisGlnIysSerProPheG 620
DB 1982 GGCAGAGGCCAAGTGAACAGAGCTGACTCGCGGCGAGCTGGCAGTGAAGAGAGCCCTTTG 2041
QY 620 lULysGlnPheIysArgArgSerCysGlnMecGlnPheGlyIysSerIleMetSerGluA 640
DB 2042 AAAAGAGTTTAAACCCAGAGCTGCCAATGGAAATTGGAGAGGACATCATGTCCAGAA 2101
QY 640 snArgSerArgGlnIysLeuGlyIysValGlySerGlnSerSerPheSerGlySerMecG 660
DB 2102 ACAGGTCAAGGAGAGCTGGGGAAAGTGGAGCAGTCAAGTCTTTTCGGGAGCATGG 2161
QY 660 lULleIleGlnValSer 665
DB 2162 AATCATTTGAGGTCTCC 2178

RESULT 14
ID ABV20833
ID ABV20833 standard; cDNA; 5145 BP.
AC ABV20833;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 20824.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-018319P.
PR 16-MAR-2000; 2000US-018962P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JF;
XX

DR WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer; useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
PS Claim 1; Page 3419; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
Alignment Scores:
Pred. No.: 0 Length: 5145
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 5 Gaps: 0
US-10-029-345A-109 (1-665) x ABV20833 (1-5145)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
DB 589 ATGCCCATGATGATTGGAATCAAAATTGACTGAGAGGTGGCTCTGCTGGAA 648
QY 21 SerGlyThrGlnIysValLeuLeuIleAspSerArgProPheValGlyTyrAsnThrSer 40
DB 649 AGTGAAGCGAAGAAAGCTGCTGTAATTGATGACCGCGCAATTTGGAAATCAATCATCC 708
QY 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerIysLeuMetIysArgArgLeuGln 60
DB 709 CACATTTTGAAGCATTATATATCACTGCTCCAAAGCTTATGAACGAAGGTGCAACAG 768
QY 61 AspIysValLeuIleThrGlnLeuIleGlnHisSerAlaIysHisIysValAspIleAsp 80
DB 769 GACAAAGTTTAAATTACAGAGCTCATCCAGCAATTCAGCGAAACATTAAGTTGACATTGAT 828
QY 81 CysSerGlnIysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 829 TGCAGTCAAGAGGTTGTGATTACATGATCAAGCTCCAAAGATGGCTCTCTCTCTTCA 888
QY 101 AspCysPheLeuThrValLeuLeuGlnIysLeuGlnIysSerPheAsnSerValHisIleu 120
DB 889 GACTGTTTTCACCTGACTTCTGGGTAAACCGAAGAGAGCTTCAACTGCTTCACTCG 948
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGlnIlyIysSer 140
DB 949 CTTCAGAGTGGGTTTGTGAGATTCTCTGTTTTCCTCGGCTCTGTAAGGAATAATCC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 1009 ACTTAGTCCCTACCTGCAATTTCTAGCTTGCTTACCTGTTGCCAAACGTGGCCAAAC 1068
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIysGlnLeuIle 180
DB 1069 CGAATTTTCCCAATCTTATCTTGGCTGCCAGAGATGCTCTCAACAGAGAGCTGAT- 1127
QY 181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProIysProAsp 200
DB 1128 GCAGCAGATGGAGTGTATGTGTAAATGCGCAAA-TACCTGTCCAAAGCTGACT 1186

QY	200	heileProgluSerHisPheLeuAaGValProValAsnaAspSerPheCysGluYsIleL	220
Db	1167	TTATCCCGAGCTCATATTCTCGCGTGGCCCTGGTAAGAAGAGCTTTGTGAGAAATTT	1246
QY	220	eUPrTriPLeuAspLysSerValAaAPheIleGluLysAlaLysAlaSerAnGlyCysV	240
Db	1247	TGCCGTGTTGGACAAATCAGTNGATTTTCATTGAGAAAGCAAAAGCTTCATGGATGTG	1306
QY	240	alLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaYrIleM	260
Db	1307	TTCTAGTGACAGTTTAGCTGGGATCTCCCGCTCCGACCACTGCTATCGCTACATCA	1366
QY	260	eLysArgIleGlnCaspMetSerLeuAspGluAlaIaYrArgPheValIleGluYsArgP	280
Db	1367	TGAAGAAGATGGACATGTCTTAAAGTAGAGCTTACAGATTTGTGTAAAGAAAAAACA	1426
QY	280	hrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTrpGluYsIleYsA	300
Db	1427	CTATATCTTCANACTTCATTTTCTGGGGCACTCTTGACTATAGAAAGAGATTAAAG	1486
QY	300	enGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGluYsPro	320
Db	1487	ACCAAGCTGGAGCATCAGGGCCAAAGAGAAACTCAAGCTGTGTGACCTGGAGAAAG	1546
QY	320	enGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProp	340
Db	1547	ATGAACCTGTCCCTCTGCTCTCAGAGGGTGGACAAAGAAAGCAGACGGCCCTCAG	1606
QY	340	roCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProAlaHisProAlaSer	360
Db	1607	CCTGGCCGAGCTCTCTACTCTCAGAGGCAGCAGGAAAGGCCCTGTGATCCCGACAG	1666
QY	360	alProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAl	380
Db	1667	TGCCAGAGCTGCCAGCGCTGCAGCCGCTGCTTGAAGACAGCCCGCTGGACAGGGC	1726
QY	380	eusSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAnLysLeuYsArgSer	400
Db	1727	TCAGTGGGCTGCACCTGTCTCCGACAGCGCTGGAAAGACAGCAATTAAGCTCAAGC	1786
QY	400	heserLeuAspLileYsSerValSerYrSerAlaSerMerAlaIleSerLeuHisGly	420
Db	1787	TCTCTGTGATACAAATACATTATTCATTTACGACAGATGGACAGATCTTACATGCT	1846
QY	420	heserSerSerGluAspAlaLeuGluYrYrLysProSerThrThrLeuAspGlyThr	440
Db	1847	TCTCTCATCAGAAATGCTTTGGAAATCTACAAACCTTCCACTACTGTGATGGAGCA	1906
QY	440	enLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSer	460
Db	1907	ACAAAGCTATGCGAGTTCTCCCTGTTACAGAACTATCGAGAGAGACTCCCGAAACA	1966
QY	460	roAspLysGluGlnLysIleProLysLeuGlnThrAlaArgProSerAspSerG	480
Db	1967	CTGAATAAAGAGAAACCCAGACATCCCAAGAACTCAGACCGCGGCTTACAGACACC	2026
QY	480	InseryArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu	500
Db	2027	AGAGCAAGCATTCGATTCGTCAGAACCAAGACAGATGGGACCGCCAGAGTCCCTTT	2086
QY	500	eusSerProLeuHisArgSerGlySerValGluAspAsnYrHisThrSerPheLeuPheG	520
Db	2087	TATCTCACTGCATCGAAGTGGAGCGTGAAGAGCAATTAACACACAGCTTCCCTTTG	2146
QY	520	ylLeuSerThrSerGlnGlnHisIleuThrLysSerAlaGlyLeuGlyLeuYsGlyTrpH	540
Db	2147	GCTTTTCCACCGCAGCAGCAGCACTCAAGAACTGTGCGCTTGAGCTTAAAGGCTGCG	2206
QY	540	IsSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpYrPheA	560
Db	2207	ACTCGGATATCTTGGCCCCCAGACCTTCAACCTTCCCTTACACGACAGCTGATTTTGG	2266
QY	560	IaThrGluSerSerHisPheYrSerAlaSerAlaIleYrGlyGlySerAlaSerYrS	580

Db	2267	CCACAGAGTCTCTACACTTCTACTCTGCTCAGCCATCTACGAGGACATGCCAGTTACT	2326
Qy	580	eraIaTySerCySeSerGlnleuPProthrCySglYaepGlnValTySerValaArga	600
Db	2337	CTGCTCTACAGTCGACGACCTGCGCACTTGCGGAGACCAAGCTTATTTCTGTGGCAGGC	2386
Qy	600	rgGlnTyPProSeSerAsPaArgAlaaspSerAArgaSerTTPHlsglGluGluSerPProbeg	620
Db	2387	GGCAGGAAGCCCAAGTGAACAGAGCTACTCTCGCGCGGAGCTGGCATGTGAAGAGAGGCCCTTTG	2446
Qy	620	IuTySglnPnPhelYsArgArgSerCtySglInmetGluPnheGluYcuUserTlleMetSerGlna	640
Db	2447	AAAAGCAGTTTAAACGACGAAGCTGCCAAATGTGAATTTGAGAGAGCATATGTCAAGA	2506
Qy	640	snArGSeSerArgGluGluLeuGluYlySerValGlySerGlnSerSerPheSerGlySermetG	660
Db	2507	ACAGGTCAACGGGAAGAGACTGGGGAATGGGAGACTGACTGTTTCCGGCAGCATGG	2566
Qy	660	IuIleIleGluValSer	665
Db	2567	AAATCATTGAGGTCTCC	2583
RESULT 15			
ABV21080	ID	ABV21080 standard; cDNA; 5145 BP.	
XX	AC	ABV21080;	
XX	DT	13-SEP-2002 (first entry)	
XX	DE	Human prostate expression marker cDNA 21071.	
XX	KM	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;	
XX	KW	pharmacogenomic marker; gene; ss.	
OS		Homo sapiens.	
XX	PN	WO200160860-A2.	
XX	PD	23-AUG-2001.	
XX	PF	20-FEB-2001; 2001WO-US0005171.	
XX	PR	17-FEB-2000; 2000US-0183319P.	
XX	PR	16-MAR-2000; 2000US-0189862P.	
XX	PR	25-MAY-2000; 2000US-0207454P.	
XX	PR	09-JUN-2000; 2000US-0211314P.	
XX	PR	18-JUL-2000; 2000US-0219007P.	
XX	PR	13-DEC-2000; 2000US-0255281P.	
PA		(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.	
XX	P1	Schlegel R, Endege WO, Monahan JB;	
XX	DR	WPI; 2001-662795/76.	
XX	PT	Novel isolated nucleic acid molecule associated with cancerous strate of	
XX	PT	prostate cells and correlating with presence of prostate cancer, useful	
XX	PT	for detecting presence of prostate cancer, stage of prostate cancer.	
XX	PS	Claim 1; Page 3481; 11750pp; English.	
CC	The invention relates to an isolated nucleic acid molecule (I) comprising		
CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the		
CC	specification or its complement. (I) is useful for: (a) assessing whether		
CC	a patient is afflicted with prostate cancer; (b) monitoring the		
CC	progression of prostate cancer in a patient; (c) assessing the efficacy		
CC	of a test compound to inhibit prostate cancer in a patient; (d) assessing		
CC	the efficacy of a therapy for inhibiting prostate cancer in a patient;		
CC	(e) selecting a composition for inhibiting prostate cancer in a patient;		
CC	(f) assessing the prostate cell carcinogenic potential of a compound; (g)		
CC	determining whether prostate cancer has metastasized in a patient; (h)		

CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP, 1408 A, 1135 C, 1253 G, 1346 T, 0 U, 3 Other;

Alignment Scores:

Pred. No.:	0	Length:	5145
Score:	472.00	Matches:	663
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	70.98%	Indels:	4
DB:	5	Gaps:	0

US-10-029-345A-109 (1-665) X ABV21080 (1-5145)

QY	MeAlHhIGlUwecTleglYThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu	20
Db	ATGGCCCAAGATGATGATTGAAACTCAAAATTGTACTGAGAGAGTGGGCTCTGCTGAA	648
QY	21 SerGlyThrGluLeuValLeuLeuIleAspSerArgProPheValGluTyraThrSer	40
Db	649 AGTGGAAACGGAAAAAGTGGCTGCTAAATTGATGGCCGGCATTTGGGAATCAAAATACATCC	708
QY	41 HisIleLeuGlnAlaIleAsnIleAsnCysserIlysMetIlyAspArgLeuGln	60
Db	709 CACATTTTGGAAAGCATTATATCAACAGCTCTCAAGCTTATGAAGCGAAGGTGCACAG	768
QY	61 AspIlyValLeuIleThrGluLeuIleGlnHisSerAlaIlysIlysValAspIleAsp	80
Db	769 GCAAAAGTGTAAATTACAGAGCTCATCCAGATTACAGGAACAATPAAAGTTGACATTGAT	828
QY	81 CysSerGlnIlyValValIlyraAspIlnSerSerGlnAspValAlaSerLeuSerSer	100
Db	829 TGCAGTCAGAAAGTGTGACTTATTCAGATCAAAAGCTCCAGAAAGTGGCTCTCTCTTCA	888
QY	101 AspCysPheLeuThrValLeuLeuGlnIlyIlyLeuGlnIlySerPheAsnSerValHisLeu	120
Db	889 GACTGTTTCTCAGCTGACTCTGGGTAACTGGAGAAAGACTTCAACTCTGTTCACTG	948
QY	121 LeuAlaGlnIlyPheAlaGlnPheSerArgCysPheProGlnIlyLeuCysGlnIlyIlySer	140
Db	949 CTTCGACAGTGGGTTTGTCTGAAGTCTCTGTGTGTTTCCCTGGCCCTGTGTGAAGGAAAAATCC	1000
QY	141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	1009 ACTCACTGCTCCACTACGCAATTTCTCAGCCTTGCTTACCTGTGCCAAACATTTGGGCCAAC	1066
QY	161 ArgIleLeuProAsnLeuTyrlLeuGlnCysGlnArgAspValLeuAsnIlyGlnLeuIle	180
Db	1069 CGAAATCTTCCAAATCTTATCTTGGCTGCCAGGAATGTCTTCAACAAGAGGCTGAT-	1127
QY	181 -GlnGlnAsnGlyIleGlyTyrlValLeuAsnAlaSerTyr-ThrCysProIlyProAsp	200
Db	1128 GCAGCAGATGGGATTTGGTTATGTATGTTAAATGCCAGAA-TACTGTCCAAAGCTGACT	1188
QY	200 HeIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnIlyIleu	220
Db	1187 TTATCCCCGAGCTCATTTCTCGGTGGCTGTGAATGACACATTTTGTGAAGAAATTT	1246
QY	220 euProTyrLeuAspIlySerValAspHeIleGlnIlysaIlysaIlysaIlysaSerAsnGlyCysV	240
Db	1247 TSCCGTGGTGGACAAATCACTAGATTTCATGGAAGCAAGAAAGCTCCAAATGATG	1306
QY	240 aIleuValHisCysLeuAlaGlnIlyIleSerArgSerAlaThrIleAlaIleAlaTyrlIem	260
Db	1307 TTCTAGTCCACGTGTTTACTGTGGATCTCCCGCTCCGCACATCCGCTATCGCTCATCATCA	1366
QY	260 eTlysaArgMetAspMetSerLeuAspGlnAlaTyrlArgPheValIlyGlnIlysaArgProT	280
Db	1367 TGAAGAGATGACATGTCTTTAATGAAGCTTTCAGATTTTGTGAAGAAAAAGACTTA	1428
QY	280 hTlIleSerProAsnPheAsnPheLeuGlnIlyLeuLeuAspTyrlGlnIlysaIlyIlysa	300

QY 660 lulellegluvalser 665
 DB 2567 AAATCATTGAGTCTCC 2583

RESULT 16

ABV26680
 ID ABV26680 standard; cDNA; 5145 BP.

XX ABV26680;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 26671.

XX Human, prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX MO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001MO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-025281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of

XX prostate cells and correlating with presence of prostate cancer, useful

XX for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 5388-5389; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

XX a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the

XX specification or its complement. (I) is useful for: (a) assessing whether

XX a patient is afflicted with prostate cancer; (b) monitoring the

XX progression of prostate cancer in a patient; (c) assessing the efficacy

XX of a test compound to inhibit prostate cancer in a patient; (d) assessing

XX the efficacy of a therapy for inhibiting prostate cancer in a patient;

XX (e) selecting a composition for inhibiting prostate cancer in a patient;

XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)

XX determining whether prostate cancer has metastasized in a patient; (h)

XX assessing the aggressiveness or indolence of prostate cancer in a patient

XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

XX Alignment Scores:

XX Pred. No.: 0

XX Score: 472.00

XX Percent Similarity: 99.40%

XX Best Local Similarity: 99.40%

XX Query Match: 70.98%

XX DB: 5

Length: 5145
 Matches: 663
 Conservative: 0
 Mismatches: 2
 Indels: 4
 Gaps: 0

US-10-029-345A-109 (1-665) x ABV26680 (1-5145)

QY 1 MetAlHieglumetiieglyThrglnIleValThrglnuArgueValAlaleuenglu 20
 DB 589 ATGGCCATGATGATGATGGAAGTCAATTTGTTACTGAGAGTTGGTGGCTCTGCTGGA 648

QY 21 SerGIYThrGIuIyValIleuIleAaspSerArgProPheValGIuYrAenThrSer 40
 DB 649 AGTGAACGGAAAAATGCTGCTAATTGATAGCCGGCCATTGTGAAATACATACATCC 708

QY 41 HisIleuGIuAlaIleasnIleAenCySerIyLeuMetIyAargArgIengIn 60
 DB 709 CACATTTTGGAGCCATTATATCACTGCTCCAGCTTATGAAAGGAGTTGCAACAG 768

QY 61 AspIyValIleuIleThrgIleuIleGIuHisSerAlaIyHisIyValAspIleAap 80
 DB 769 GACAAAGTGTATTATACAGAGCTCATCCAGCATTCAGCAAAACATAGAGTTGACATTGAT 828

QY 81 CysSerGIuIyValIyValIyTyrAspGINserSerGIaAspValIAserIeuserSer 100
 DB 829 TGCAGTCAGAGGTGTAGTTTACATCAAGCTCCCAAGATGTCTCTCTCTCTTCA 868

QY 101 AspCyPheIeuThrValIleuengIyIySLeuGIuIySerPheAnsSerValHisIleu 120
 DB 889 GACTGTTTCTACCTGCTACTCTGCGGTAAACTGGAGAGAGCTTCAACTCTGTTCACCTG 948

QY 121 LeuAlaGIyGIyPheAlaGIuPheSerArgCyPheProGIyLeuCySGIuGIyIySer 140
 DB 949 CTTCAGGTGGGTGTGCTGAGTCTCTCGTTGTTCTGCTCTGTGAGAGAAATCC 1008

QY 141 ThrIeuValProThrCysIleSerGINProCySLeuProValIAsenIleGIYProThr 160
 DB 1009 ACTTATGCTCCCTACCTGCAATTTCTCAGCCTGTGCTTACCTGTGCAACATTTGGCCAAAC 1068

QY 161 ArgIleuProAnsIeuThrIeuGIyCySGInArgAspValIleuAnsIyGIuIle 180
 DB 1069 CGAATTCCTCCCAATTTTATCTGCTGCTCCAGCAGAGAGTCTCTCAACAGAGCTGAT- 1127

QY 181 -GInGIuAnGIyIleGIyIyValIleuAnsIAserIy-ThrCySProIyProAap 200
 DB 1128 GCAGCAGATGGAGTGTATGTATGTAAATGCCAGCAA-TACCTGTCCAAAGCTGACT 1186

QY 200 helIeProGIuSerHisPheIeuArgValProValAnsPhePheCySGIuIyIleI 220
 DB 1187 TTATCCCGAGTCTCATTTCTGCGTGTGCTGTGAATACACCTTTGTGAAGAAATTT 1246

QY 220 euProTrIeuAnsPlySerValAspPheIleGIuIySAlaIySAserAnGIyCyv 240
 DB 1247 TGCCTGTGTGAGCAATCATGATTTATTGAGAAAGCAAAAGCTCCATGATGTG 1306

QY 240 alleuValHisCySLeuAlaGIyIleSerArgSerAlaThrIleAlaIleAlaIyIle 260
 DB 1307 TTCTAGTGCACGTGTTAGCTGGGATCTCCGCTCCGCAACATCGCTATCGCTACATCA 1366

QY 260 elIyEAArgMetAspMetSerIeuAspGIuAlaIyArgPheValIyGIuIyArgProt 280
 DB 1367 TGAAGAGATGACATGCTTTTATGAGAGCTTACAGATTGTGAAAGAAAGAACCTTA 1426

QY 280 hIleSerProAnsPheAnsPheIeuGIuIleuAnsPlyrGIuIySValIleIyA 300
 DB 1427 CTATATCTCAAACTTCAATTTCTGGGCCAACTCTGCACTATGAGAAAGATTAAG 1486

QY 300 enGIuThrGIaIAserGIyProIySerIySLeuIySLeuHisIleuGIuIyProA 320
 DB 1487 ACCAGACTGGAGCATGAGGCCCAAGAGCAAACTCAAGCTGTGCACTGGAGAGACCA 1546

QY 320 enGIuProValProAlaValIserGIuGIyGIuIySergIuThrProIeuSerProp 340
 DB 1547 ATGAACCTGTCTGCTGTCTGCAAGGGGTGAGCAAGAAAGCGAGCCCTCAGTCCAC 1606

QY 340 rOCyAlaAspSerAlaThrSerGIuAlaIaGIuIyGIuIyProValHisProIaSerV 360
 DB 1607 CCTGTGCGACTGTGTACTTCAAGGCGAGCGAGACAAAGGCCCGGTGATCCGCCAGCG 1666

QY 360 aIProSerValProSerValGIuProSerIeuGIuIyAspSerProIeuValGIuAla 380
 DB 1667 TGCCAGCGCTGCGCAGCGGTGAGCGCGTGTAGAGAGACAGCCCGCTGTGACAGGCC 1726

```
QY 380 euserGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAnlyLeuLyAspSerP 400
    |||||
DB 1727 TCAGTGGGGTGCACCTGTCGGAGACAGGCTGGAAACAGCAATAGCTCAAGCGTTCT 1786
QY 400 hSerLeuAspIleLeuSerValSerTyrSerAlaSerMetAlaAspLeuHisGlyP 420
    |||||
DB 1787 TCTCTGATATCAAAATCAATGATTTCAATATTCAAGCCGCACTGGCAGATCTTCAACATGCT 1846
QY 420 hSerSerSerGluAspAlaLeuGluTyrTyrLeuProSerThrThrLeuAspGlyThrA 440
    |||||
DB 1847 TTTCTCATCAAGAGATGCTTTGGATTAATCAAACTTCCACTACTCTGGATGGACCA 1906
QY 440 snlyLeuCySGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerP 460
    |||||
DB 1907 ACAAGTATGCCAGATTCTCCCTGTTCAAGAACTATCGAAGAGACTCCGAAACAGAGC 1966
QY 460 roAspLySGluGluAlaSerIleProLyLeuGlnThrAlaArgProSerAspSerG 480
    |||||
DB 1967 CTGATTAAGAGAGAGAGCAGCATCCCAAGAGCTGAGAGCCGCAAGCCTTCAACAGCC 2026
QY 480 InSerLyAspLeuHisSerValArgThrSerSerSerGlyThrAlaArgSerLeuL 500
    |||||
DB 2027 AAGAGAGGATTCGATTCGTCAGAACAGACAGAGAGAGCCGCAAGAGTCCCTTT 2086
QY 500 euserProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheG 520
    |||||
DB 2087 TATCTCCATCGATCGAAGTGGAGGCTGGAGGACATTAACACACACACTTCTTTTG 2146
QY 520 llyLeuSerThrSerGlnGlnHisLeuThrLySerAlaGlyLeuGlyLeuLyGlyTPH 540
    |||||
DB 2147 GCTTTTCCACAGCCAGCAGACCTCTCAACAAAGTCTGCTGGCCCTTAAAGGCTGAGC 2206
QY 540 iSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTPYrPheA 560
    |||||
DB 2207 ACTCGGATATCTTGGCCCCCAGACCTCTAACCCCTTCTGACACAGACTGTATTTTG 2266
QY 560 laThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrS 580
    |||||
DB 2267 CCACAGAGTCTCAACACTTCTACTGCTGCTCAGCCATCTAAGAGGAGAGTGCAGTTACT 2326
QY 580 eAlaIAspSerCySGlnLeuProThrCySGlyAspGlnValTyrSerValArgArgA 600
    |||||
DB 2327 CTGCTTACAGCTGCAGCCAGCCAGCTGCGAGAACCAAACTTAATTTGTCGCAAGCC 2386
QY 600 rGlnLySerProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSerProPheG 620
    |||||
DB 2387 GGCAGAGCCAGACTGACAGAGCTGACTCCGCGCGAGAGCTGGCATGAGAGAGCCCTTTG 2446
QY 620 llyLeuGlnPheLyAspArgSerCySGlnMetGluPheGlyGluSerIleMetSerGluA 640
    |||||
DB 2447 AAAAGCAGTTTAAACGACAGAGCTGCCAATGGAATTTGGAAGAGCATCATGTCAAGA 2506
QY 640 snArgSerArgGluGluLeuGlyLyValGlySerGlnSerSerPheSerGlySerMetG 660
    |||||
DB 2507 ACAGGTCAAGGAGAGGCTGGGAAAAGTGGCAGTCACTAGCTTTTCGGCAGACATGG 2566
QY 660 lylleIleGluValSer 665
    |||||
DB 2567 AAATCATTAAGGTCTCC 2583
    |||||

RESULT 17
ABV20978
ID ABV20978 standard; cDNA; 5145 BP.
XX
AC ABV20978;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 20969.
XX
KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
```

```
OS Homo sapiens.
XX
PN MO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-018313P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1, Page 3451; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
XX
Alignment Scores:
Pred. No.: 0 Length: 5145
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 5 Gaps: 0
XX
US-10-029-345A-109 (1-665) x ABV20978 (1-5145)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
    |||||
DB 589 ATGGCCCAATGAGATGATTTGAACTCAATTTTCTGAGAGAGTGGTGGCTCTGCGTGA 648
QY 21 SerGlyThrGluLyValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
    |||||
DB 649 AGTGAACGGAANAAGTCTCTGTAATATAGCCGCGCATTTGGATTAACATATCATCC 708
QY 41 HisIleLeuGluAlaIleAsnIleAsnCySGlnSerLyLeuMetLyAspArgLeuGln 60
    |||||
DB 709 CACATTTTGAAGCATTAATATCAATCACTCTCAAGCTTATGAAGCGAAAGTTGCAACAG 768
QY 61 AspLyValLeuIleThrGluLeuIleGlnHisSerAlaLyHisLyValAspIleAsp 80
    |||||
DB 769 GACAAAGTTAATTAACAGACTCATCCAGATTCCAGAAACATTAAGTTGACATTGAT 828
QY 81 CysSerGluLyValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
    |||||
DB 829 TGCAGTCAGAGGTTGTGATTACATCAAGAGCTCCCAAGATGTTGCTCTCTCTTCA 888
QY 101 AspCySPheLeuThrValLeuLeuGlyLyLeuGluLySerPheAsnSerValHisLeu 120
    |||||
```

```

Db      889  GACTGTTTCTACCTGTTCTGGGTTAACTGGAGAAGAGCTTCAACTCTGTTCACTCG 948
Qy      121  LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlyLeuSer 140
Db      949  CTTGCGAGGTGGGTTTCTGAGTTCTCTCGTTGTTTCTCGGCTCTGTAAGAGAAATTC 1008
Qy      141  ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db      1009  ACTTAAGTCCCTACCTGCACTTCTCAGCCTTGCTTAAGCTGTTGCCAACAATGGGCAACC 1068
Qy      161  ArgIleLeuProAsnLeuTyrlleuGlyCysGlnArgAspValLeuAsnIleGlyLeuIle 180
Db      1069  CGAATCTTCCCAATCTTTATCTGCTGCGACGAGATGCTCTCAACAAGAGCTGAT- 1127
Qy      181  -GlnGlnAsnGlyIleGlyTyrlleuAlaLeuAsnIleSerTyrlleuProCysProAsp 200
Db      1128  GCAGCAGAAATGGATGGTTATGTGTTAAATGCCAGCAA-TACCTGTCCAAAGCTTGACT 1186
Qy      200  HeIIeProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleL 220
Db      1187  TTATCCCGCAGATCTCAATTCCTGCGTGTGCTGTGAATGACAGCTTTGTGAGAAATTT 1246
Qy      220  euProTrpLeuAspIleSerValAspPheIleGluIleValAlaValAsnIleCysArg 240
Db      1247  TGCCTGCTTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCAAATGATGTG 1306
Qy      240  AlIleuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlleu 260
Db      1307  TTCTACTGACATGTGTAGTGGGATCTCCGCTCCGCCACATCGCTATCGCTCATCA 1366
Qy      260  eIIyAsArgMetAspMetSerLeuAspGluAlaTyrlleuPheValIleValGluValArgPro 280
Db      1367  TGAAGGAGATGACATGCTTTAGATGAGCTTACAGATTGTGAAAGAAAAAGACCTTA 1426
Qy      280  hrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrlleuValIleValAla 300
Db      1427  CTRATCTCCAACTTCAATTTCTGCGCAACTCTGCACTATGAGAAAGATTAAGA 1486
Qy      300  enGlnThrGlyAlaSerGlyProIleSerIleLeuIleValIleLeuGluIleValPro 320
Db      1487  ACCAGACTGGAGATCAGGGCCAAAGCAAACTCAAGCTGCTGACCTGAGAAACCAA 1546
Qy      320  enGluProValProAlaValAsnGlyGlyGlnIleValSerGluThrProLeuSerPro 340
Db      1547  ATGAACCTGCTCCTGCTGCTCAGAGGAGGAGCAAGAAAGCGAGAGCCCTTAAGTCCAC 1606
Qy      340  roCysAlaAspSerAlaThrSerGluAlaAlaGlyIleArgProValHisProAlaSerVal 360
Db      1607  CCTGTCCCACTGCTGCTACTCTCAGAGGACAGAGCAAAAGCCCGTGCATCCGCGACGCG 1666
Qy      360  alProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAla 380
Db      1667  TGCCCAAGCGGCGCCAGCGTGCAGCCGTCGCTTTAGAGGACAGCCGCTGTACAGCGCG 1726
Qy      380  euSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIleValIleValArgSer 400
Db      1727  TCAGTGGGCTGCACCTGTGCGAGACAGGCTGGAACAGCAATTAAGCTCAACGCTTCT 1786
Qy      400  heSerLeuAspIleLeuSerValSerTyrlleuAlaSerValAlaIleSerLeuHisGly 420
Db      1787  TCTCTCTGATATCAATCAATCAATTCATTAATTCACGACATGAGCATCTTCAATGAGCT 1846
Qy      420  heSerSerSerGluAspAlaLeuGluTyrlleuValProSerThrThrLeuAspGlyThr 440
Db      1847  TCTCTCATCAGAAAGATGCTTTGGAATATCAAAACCTTCACTACCTGATGAGGACCA 1906
Qy      440  enIleValCysGlnPheSerProValGlnIleLeuSerGluGlnThrProGluThrSer 460
Db      1907  ACAAGCTATGCGCAGTCTCCCTGTTCAAGAACTATCGAGCAGACTCCCGAAACCAAGTC 1966
Qy      460  roAspIleValGluAlaSerIleProValIleLeuGlnThrAlaArgProSerAspSer 480

```

```

Db      1967  CTGATTAAGAGGAAACCCAGCATCCCAAGAAAGCTGCAGACCCGACCTTCAGACGCC 2026
Qy      480  InsIleValArgLeuHisIleSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
Db      2027  AAGACAGAGGATGATTCGTGTCAGAAACGAGCAGTGGCACCCGCGCAGAGGCTCTTT 2086
Qy      500  euSerProLeuHisArgSerGlySerValGluAspAsnTyrlleuThrSerPheLeuPhe 520
Db      2087  TATCTCACTGATCGAAGTGGAGCGTGGAGACATTAACACACCACTTCTCTTTTG 2146
Qy      520  ILeuSerThrSerGlnGlnHisLeuThrIleSerAlaGlyLeuGlyLeuValGlyTyrlleu 540
Db      2147  GCCTTCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2206
Qy      540  IeSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrlleuPhe 560
Db      2207  ACTCGATATCTTGGCCCCCGCCAGACCTTAACCTTCCCTGACACGACAGCAGCTGATTTTG 2266
Qy      560  IaThrGluSerSerHisPheTyrlleuSerAlaSerAlaIleTyrlleuGlySerAlaSerTyrlleu 580
Db      2267  CACAGAGTCTCAACATCTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2326
Qy      580  exAlaTyrlleuSerCysSerGlnLeuProThrCysGlyAspGluValTyrlleuValArg 600
Db      2327  CTGCTTACAGCTGACAGCAGCTGCCCATTTGCGAGAACCAAGTCTATCTGTGCGCAGGC 2386
Qy      600  rGlnIleProSerSerAspAlaAspSerArgArgSerThrIleGluGluSerProPhe 620
Db      2387  GGCAGAGCAGCAAGTGCAGAGCTGACTCGCGCGAGCTGGAGATGAAAGAGAGCCCTTTG 2446
Qy      620  IuIleGlnPheIleValArgSerCysGlnMetGluPheGlyGluSerIleMetSerGlu 640
Db      2447  AAAAGCAGTTTAAAGCAGAACTGCTCAATGGAATTTGGAAGAGCATCATGTCAGAA 2506
Qy      640  snArgSerArgIleGluLeuGlyIleValGlySerGlnSerSerPheSerGlySerMet 660
Db      2507  ACAGGTCAGGGAGAAGCTGGGGAAGTGGCAGTCACTGCTTTTGGGCGCAGATCG 2566
Qy      660  IuIleIleGluValSer 665
Db      2567  AAATCATGAGGCTTCC 2583

RESULT 18
ABV21092
ID ABV21092 standard; cDNA, 5145 BP.
AC ABV21092;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 21083.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-018319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JB;

```


XX WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 3485; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (1) comprising a nucleotide sequence given in Tables 1-9 (ABV0010-ABV6213) of the CC specification or its complement. (1) is useful for: (a) assessing whether CC a patient is afflicted with prostate cancer; (b) monitoring the CC progression of prostate cancer in a patient; (c) assessing the efficacy CC of a test compound to inhibit prostate cancer in a patient; (d) assessing CC the efficacy of a therapy for inhibiting prostate cancer in a patient; CC (e) selecting a composition for inhibiting prostate cancer in a patient; CC (f) assessing the prostate cell carcinogenic potential of a compound; (g) CC determining whether prostate cancer has metastasized in a patient; (h) CC assessing the aggressiveness or indolence of prostate cancer in a patient CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 5145 BP, 1408 A, 1135 C, 1253 G, 1346 T, 0 U, 3 Other;

Alignment Scores:

Pred. No.:	0	Length:	5145
Score:	472.00	Matches:	663
Percent Similarity:	99.408	Conservative:	0
Best Local Similarity:	99.408	Mismatches:	2
Query Match:	70.988	Indels:	4
DB:	5	Gaps:	0

US-10-029-345A-109 (1-665) x ABV21092 (1-5145)

QY	1	MetAlaHisGluMetIleGlyThrGlnIleValIThrGluArgLeuValAlaLeuLeuGlu	20
DB	589	ATGGCCCAAGAGATGATGGAATCTCAATTTGTTACTGAGAGAGGTGGTGGCTGGCGAA	648
QY	21	SerGlyThrGluLeuValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer	40
DB	649	AGTGGACCGAAGAAAGTGGCTGCTAATGTAGACCGGCGCATTTGTGAAATACAAATCATCC	708
QY	41	HisIleLeuGluAlaIleAsnIleAsnGlySerLeuMetLeuArgArgLeuGlnGln	60
DB	709	CACATTTTGGAGCCATTATATCACTGCTCCAGCTTATGAGGAGAGGTGGCAACG	768
QY	61	AspLeuValLeuIleThrGluLeuIleGlnHisSerAlaLeuHisIleValAspIleAsp	80
DB	769	GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTCAGCGAACAATAGAGTTGACATTTGAT	828
QY	81	CysSerGlnLeuValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer	100
DB	829	TGCAGTCAGAGAGTTGTACTTATCAATCAAGCTCCCAAGATGTGGCTCTCTCTTCA	888
QY	101	AspCysPheLeuThrValLeuLeuGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle	120
DB	889	GACTGTTTCTCAGTGTACTTCTGGGTAAACTGGAGAGAGCTTCAACTCTGTTCACCTG	948
QY	121	LeuAlaGlyLeuPheAlaGluPheSerArgCysPheProGlyLeuGlnGlnIleGlnIle	140
DB	949	CTTGGAGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1008
QY	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
DB	1009	ACTTAGTCCCTACTGCAATTCAGCTTGTCTTACCTGTGGCCAACTGGGCAACC	1068
QY	161	ArgIleLeuProAsnLeuTyrLeuGlnCysGlnArgAspValLeuAsnIleGlnIle	180
DB	1069	CGAATTCCTCCCAATCTTATCTTGGCTGCCAGCAGAGATGCTCTCAACAGAGAGCTGAT-	1127
QY	181	GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProIysProAsp	200
DB	1128	GCAGCAGATGGATGGTGTATGTATGTTAATGCCAGCAA-TACGTGTCCAAAGCTGACT	1186

QY	200	helleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlnIleLeu	220
DB	1187	TTATCCCGAGCTCATTTCTGGCTGTGCTGTAAGACGCTTTGTGGAATAATTT	1246
QY	220	eupProTyrLeuAspIysSerValAspPheIleGlnIleGlnIleGlnIleGlnIleGlnIle	240
DB	1247	TGCCGTGGTGGACAAATCAGTAAATTCATTGAGAAAGCAAAAGCTCCCATAGGATGTG	1306
QY	240	alleuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAlaIle	260
DB	1307	TTCTAGTCACTGTTTACTGCTGGATCTCCCTCCCTCCGACCATGCTTCCCTACATCA	1366
QY	260	eLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGlnIleGlnIleGlnIle	280
DB	1367	TGAAGAGATGACATGCTTATAGTAGCTTACAGATTTGTGAAAGAAAAAGACCTTA	1426
QY	280	hrlLeuSerProAsnPheAsnPheLeuGlnIleLeuAspTyrGlnIleGlnIleGlnIleGlnIle	300
DB	1427	CTATATCTCCAAACTTCAATTTCTGGGCCAACTCTGACTATGAGAAAGAAATTAGA	1486
QY	300	enGlnThrGlyAlaSerGlyProIysSerIleuLysLeuLeuHisLeuGlnIleGlnIleGlnIle	320
DB	1487	ACCGAGCTGGAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGACCTGGAGAGGCCAA	1546
QY	320	enGluProValProAlaValSerGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle	340
DB	1547	ATGAACTCTCTCTGCTGTCTCAAGAGGTGACAGAAAGCGAGAGCCCTCAGTCCAC	1606
QY	340	roCysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSerV	360
DB	1607	CTGTGGCGGACTGTCTATCTCTCAGAGGCGAGGCAAAAGCCCGTGTATCCCGCAGCG	1666
QY	360	alProSerValProSerValGlnProSerLeuLeuGlnIleAspSerProLeuValGlnAlaL	380
DB	1667	TGCCAGGCTGCCAGCGCTGAGCGCTGCTGTTAGAGACAGCCCGTGTATACAGGCGC	1726
QY	380	eusSerGlyLeuHisLeuSerAlaAspArgLeuGlnIleAspSerAsnIleGlnIleGlnIleGlnIle	400
DB	1727	TGAGTGGCTGACCTGCTGCCAGACAGCTGGAAGACAGCAATAGCTCAAGAGCTTCT	1786
QY	400	heserLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyP	420
DB	1787	TCTCTGGAATATCAATCAATTCATTATTCAGCAGATGGAGCATCTTATCATGCT	1846
QY	420	heserSerSerGlnAspAlaLeuGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle	440
DB	1847	TCTCTCATCAGAAAGATGCTTTGGAATATCAAAACCTTCCATCACTGTGATGGACCA	1906
QY	440	snLysLeuCysGlnPheSerProValGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle	460
DB	1907	ACAAAGCTATGCCAGTCTTCTCCCTGTTACAGAACTATGGAGCAACTCCGAAACCAAGTC	1966
QY	460	roAspLysGlnGlnAlaSerIleProLysIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle	480
DB	1967	CTGATTAAG	2026
QY	480	InseryAspGlnLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuL	500
DB	2027	AGACCAAGCGATTCATTCGCTCAGAACACAGCAGACAGTGGACCGCCAGAGGTCCCTTT	2086
QY	500	eusSerProLeuHisArgSerGlySerValGlnIleAspAsnTyrHisThrSerPheLeuPheG	520
DB	2087	TATCTCACTGCAATCAAG	2146
QY	520	IleuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlnIleGlnIleGlnIleGlnIle	540
DB	2147	GCCTTCCACAGCAGCAG	2206
QY	540	isSerAspIleLeuAlaProGlnIleThrSerThrProSerLeuThrSerSerTyrPheA	560
DB	2207	ACTGGAATATCTTGGCCCGCCAGACCTTACCCCTTCCCTGACAGCAGCTGTGATTTTG	2266

QY 560 laThiGluSerSerHisPheTySerAlaSerAlaIleTyGlyGlySerAlaSerTyrS 580
DB 2267 CCACAGAGCTCTCACACTTCTACTTGCCTCAGCCATCTTACGAGCAGTCCAGTTACT 2326
QY 580 eArIaTySerCySerGlnLeuProThCyGlyAlaSpGlnValTySerValArgArgA 600
DB 2327 CTGGCTTACAGCTGCACACAGCTGCCCATCTTGCAGAACCAAGCTATTCTGTGCGCAGGC 2386
QY 600 rGgInTySProSeArPaRgAlaAspSerArgrgSerTPhIsgIuGluSerProPhg 620
DB 2387 GGCAGAACCCAACTGACAGAGCTGACTCGCGCGAGAGCTGGCATGAAAGAGGCCCTTTGG 2446
QY 620 JuTySgInPheTySArGrSerCySgIuMeGluPheGlyGluSerIleMeSerGluA 640
DB 2447 AAAGCAGTTTAAACCAAGAGCTGCCAAATGGAATTGGAGAGACATCATCTCGAGA 2506
QY 640 sArGrSerArgrGluGluLeuGlyLySValGlySerGlnSerSerPheSerGlySerMeG 660
DB 2507 ACAGGTCCAGCGGAGAGAGCTGGGGAAGTGGGAGTCACTAGCTTTCGGGAGCATGG 2566
QY 660 JuIleIleGluValSer 665
DB 2567 AAATCATTAAGAGCTCC 2583
RESULT 19
ABV21312
ID ABV21312 standard; cDNA; 5145 BP.
XX ABV21312;
AC
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 21303.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN MO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001, 2001WO-US005171.
XX
PR 17-FEB-2000, 2000US-018319P.
PR 16-MAR-2000, 2000US-0189862P.
PR 25-MAY-2000, 2000US-0207454P.
PR 09-JUN-2000, 2000US-0211314P.
PR 18-JUL-2000, 2000US-0219007P.
PR 13-DEC-2000, 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JB;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX
PS Claim 1; Page 3539; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP, 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
Alignment Scores:
Pred. No.: 0 Length: 5145
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 5 Gaps: 0
US-10-029-345a-109 (1-665) x ABV21312 (1-5145)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 589 ATGGCCCATGAGATGATTTGAACTCAATTTTACTGAGAGGTGGCTGTGAGAA 648
QY 21 SerGlyThrGluLySValLeuLeuIleAspSerArgrgProPheValGluTyRhenThSer 40
DB 649 AGTGAACGGAAGAAAGTGTCTTAATGATGACCGGCAATTTGTGAATACATACATCC 708
QY 41 HisIleLeuGluAlaIleAsnIleAsnCySerIleLySLeuMetIleArgIleGluGln 60
DB 709 CACATTTTGGAGCCATTATATCAATCTCTCCAGCTTATGAGGAAAGTTGGCAACG 768
QY 61 AspLySValIleuIleThrGluLeuIleGlnHisSerAlaLySValHisLyValAspIleAsp 80
DB 769 GACAAAGTGTATTAACAGAGCTCATCCAGATTGAGCAAACTAATGAGTTGACATTGAT 828
QY 81 CySerGlnLySValValValTyRAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 829 TGCAGTCAGAAAGTTGATTTTACATCAATCAAACTCCCAAGATGTGCTCTCTCTTGA 888
QY 101 AspCyPheLeuThrValLeuLeuGlyLySLeuGluLySLeuSerPheAsnSerValHisLeu 120
DB 889 GACTGTTTCTACCTGTACTTCTGGTAACTCGAGAGAGCTTCAACTCTGTCCCTG 948
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCyPheProGlyLeuCySgIuGlyLySLeu 140
DB 949 CTTCAGAGTGGGTTGTGAGTTCTCTCGTTTCCCTGGGCTCTGTGAAAGAAATCC 1008
QY 141 ThrLeuValProThCyAlleSerGlnProCySLeuProValAlaAsnIleGlyProThr 160
DB 1009 ACTTACGTCCCTACCTGCAATTCCTCAGCCTTCTTACCTGTGCCAACTTGGGCCAAC 1068
QY 161 ArgIleLeuProAsnLeuTyRLeuGlyCySgInArgAspValLeuAsnLySgIuLeuIle 180
DB 1069 CGAATTTCTCCAAATCTTTATCTTGCTGCGCAGCGAGATGTCTCAACAAAGAGCTGAT- 1127
QY 181 -GlnGlnAsnGlyIleGlyTyRValLeuAsnAlaSerTyR-ThrCySProLySProAsp 200
DB 1128 GCAGCAGATGGGATTTGGTTATGTGTAAATCCAGCAA-TACGTGTCCAAAGCCGACT 1186
QY 200 heIleProGluSerHisPheLeuArgValProValAsnAspSerPheCySgIuLySLeu 220
DB 1187 TTATCCCCAGAGCTCATTTCTCGCGTGCCTGTGATACACAGCTTTGTGAAATTT 1246
QY 220 euProTyRLeuAspLySLeuValAspPheIleGluLySValAlaLySLeuAsnGlyCySv 240
DB 1247 TGCCTGTGTGACAAATCATGATTTTCATTGAGAAACAAAGCCCTCAATGATGTG 1306
QY 240 alleuValHisCySLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyRilem 260
DB 1307 TTCTAGTGACATTTTAACTGGGATCTCCGCTCCGCAACATCGCTATCGCTACATCA 1366
QY 260 eLySArgMetAspMetSerLeuAspGluAlaTyRArgPheValLySgIuLySArgPro 280
DB 1367 TGAAGAGGTGACATGTCTTTAGATGAAGCTTACAGATTGTGTGAAGAAAAAGCTTA 1426
QY 280 hrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyRgIuLySlySleLyS 300

```

1427 CTATATCTCCAACTTCAATTTTCTGGCCCACTCCCTGACCTATGAGAGAAATTTAAGA 1486
QY 300 snglnthrGlyAlaSerGlyProLySerLySleuLySleuLH.sleuGluLyProA 320
Db 1487 ACCAGACTGAGAGATCAAGGGCCAAAGACCAACTCAAGCTGCTGACCTTGAGAAACCA 1546
QY 320 sngluProValProAlaValSerGluGlyGlyGlnLySerGluLuthrProLeuSerProP 340
Db 1547 ATGAACCTGCTCCCTGCTGTCTCAGAGGGTGAACAGAAAGCGAGAGCCCTCAAGTCAC 1606
QY 340 roCySaIaapSerAlaThrSerGluAlaAlaGlyInaPProValHIsProAlaSerY 360
Db 1607 CTTGTGGCCGACTCTGCTACTCAAGAGCAGACAGAAAGCCCGTGACATCCGCCACAGG 1666
QY 360 aIProSerValProSerValGlnProSerLeuLugluuapSerProLeuValGlnAla 380
Db 1667 TGCCCAAGCCGTCAGCGGTGCACCGCTGCTTTAGAGGACAGCCCGCTGTGTCAGAGCCG 1726
QY 380 euSerGlyLeuHIsleuSerAlaAPArghLeuGluuapSerAenLySleuLyArghSerP 400
Db 1727 TCAGTGGGCTGCACCTGTCCGCAAGAGGCTGGAGAGCAGAAATAGCTCAAGCGTTCC 1786
QY 400 heSerLeuapIleuLySerSerValSerTySerAlaSerMetalaAlaSerLeuHIsGlyP 420
Db 1787 TCTCTCTGATATCAATCAATGATTTCATATTCAAGCCAGCATGGCAGCATCTTACATGSGCT 1846
QY 420 heSerSerSerGluuapAlaLeuGluTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 440
Db 1847 TCTCTCATCAGAAAGATGCTTGGAAATACCAAACTTCCATCTCTGATGGAGCA 1906
QY 440 snLySleuCySgInPheSerProValGlnGluSerGluGlnThrProGluLuthrSerP 460
Db 1907 ACAAGCTATGCCAGTTCTCCCTGTTCAAGAACTATCGAGCAGACTCCCGAAACCAAGTC 1966
QY 460 roApLySgLuGluAlaSerTLeProLySleuGlnThrAlaArghProSerArghSerP 480
Db 1967 CTGATAGAGAGAGAGCAGACATCCCAAGAGCTGAGACCGCAGGCTTTCAGACAGCC 2026
QY 480 InSerTyArghLeuHIsSerValArghThSerSerSerGlyThrAlaGlnArghSerLeu 500
Db 2027 AGAGCAGAGATGCTTGGTGTGAGAACAGCAGCGATGACCGCCAGAGGTCCCTTT 2086
QY 500 euSerProLeuHIsArghSerGlySerValGluuapAntTyHIsThrSerPheLeuapheG 520
Db 2087 TATCTCACTGCATGCAAGTGGAGGAGGTGAGAGCAATTAACAACACAGCTTCTTTTCG 2146
QY 520 lYLeuSerThrSerGlnGlnHIsleuThrLySerAlaGlyLeuGlyLeuLyGlyTyTyTy 540
Db 2147 GCCTTTCACACAGCCAGCAGACCTCAAGAGTGTGCTGGCTTGAAGGGCTGGC 2206
QY 540 iSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerTyrTyTyTyTyTy 560
Db 2207 ACTCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCGCTGGATTTTTCG 2266
QY 560 lAThrGluSerSerHIsPheTyTySerAlaSerAlaIleTyTyGlyLySerAlaSerTyTy 580
Db 2267 CCACAGAGCTCCACACACTTCTACTCTGCTGCTGCAATCTACGAGGACGATGCAATTA 2326
QY 580 eAlaTySerCySerGlnLeuProThrCyGlyVaSPGlnValTyTySerValArghArgh 600
Db 2327 CTGCCTACAGCTGACAGCCAGCTGCCACTTGGCGAGACCAAGTCTATCTGTCGCGCAGGC 2386
QY 600 rGgIlySerProSerArghAlaapSerArghArghSerTyHIsGluLuthrSerProheG 620
Db 2387 GCGAGAGCAGAGTACAGAGTGACTGCGCGGAGCTGGAGTGAAGAGAGGCCCTTTTCG 2446
QY 620 lLySgInPheLyArghArghSerCySgInMeGlnPheGlyGluSerTLeMetSerGlu 640
Db 2447 AAAACAGATTTAAAGCAGAGAGCTCCAAATGGAATTTGGAGAGAGCATCATGTCAGGA 2506
QY 640 snArghSerArghGluGluLeuGlyLyValGlySerGlnSerSerPheSerGlySerMetG 660
```

```

Db 2507 ACAGGTCACGGGAGAGAGCTGGGGAAGTGGGCACTGACTTACTTTTCGGCAGACATGG 2566
QY 660 lAllelIleGluValSer 665
Db 2567 AATCATTTAGAGGTCTCC 2583
RESULT 20
ABV21316
ID ABV21316 standard; cDNA, 5145 BP.
XX
XX ABV21316;
AC
XX
XX 13-SEP-2002 (first entry)
DT
XX
XX Human prostate expression marker CDNA 21307.
DE
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200160860-A2.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 20-FEB-2001; 2001WO-US005171.
PF
XX
XX 17-FEB-2000; 2000US-0183319P.
PR
XX 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUN-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX
XX Schlegel R, Endege WO, Monahan JB;
PI
XX
XX WPI: 2001-662795/76.
DR
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PI for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 3540; 11750pp; English.
PS
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
Alignment Scores:
Pred. No.: 0 Length: 5145
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 5 Gaps: 0
US-10-029-345A-109 (1-665) x ABV21316 (1-5145)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArghLeuValAlaLeuLeuGlu 20
```

Db 589 ATGGCCCATGATGATGGAACTCAATTGTTACTGAGAGTGGCTGCTGCTGAA 648
 QY 21 SerGIYThGluYValLeuLeuIleAspSerArgProPheValGluYrAntHrSer 40
 Db 649 AGTGAACGGAAAAAGGCTGCTAAATTGATAGCCGGCCATTGTGGAAATACATACATCC 708
 QY 41 HsAlleuGluIuAlleuAnIleAnCySerLeuLeuMetLysArgArgLeuGlnGln 60
 Db 709 CACATTTTGGAGCCATTATATCACTGCTCAACCTTATGAGGAGAGTTGGCAACAG 768
 QY 61 AspLysValLeuIleThrGluLeuIleGlnHsSerAlaYehIshLysValAspIleAsp 80
 Db 769 GCAAAAGTGTATTCAGAGCTCATCCAGCATTCAGCCAAACATAGAGTTGACATTGAT 828
 QY 81 CysSerGluYValValIleYrAspGlnSerSerGlnAspValIleSerLeuSerSer 100
 Db 829 TGGAGTCAGAAAGGTGTAGTTTACGATCAAAAGCTCCCAAGATGTCTCTCTCTTCA 888
 QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluYSerPheAnSerValHsIleu 120
 Db 889 GACTGTTTCTCCTGACTTCTGGGTAACTGGAAAGAGCTTCAACTCTGTCACCTG 948
 QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlyLysSer 140
 Db 949 CTTCAGAGTGGGTTTCTGCTGAGTCTCTCGTTGTTTCCCTGGCTCTGTGAAGAAATCC 1008
 QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAnIleGlyProThr 160
 Db 1009 ACTCTAGTCCCTTACCTGCAATTTCTCAGCTTCTGCTTACCTGTCACAACTGGGCCAACCC 1068
 QY 161 ArgIleLeuProAnLeuThrLeuGlyCysGlnArgAspValLeuAnIleGlyLeuIle 180
 Db 1069 CGAATCTTCCCAATCTTTATCTCTGGCTGCCAGCAGAGTCTCTCAACAGAGCTGAT- 1127
 QY 181 -GlnGlnAnGlyIleGlyYrValLeuAnIleAspYr-ThrCysProLysProAsp 200
 Db 1128 GGAGCAGATGGGATTGTTATGTATTAATGCCACAA-TACCTGTCCAAAGCCGACT 1186
 QY 200 HeIleProGluSerHsPheLeuArgValProValAnAspSerPheCysGluYrIleu 220
 Db 1187 TTATCCCGCAGTCTCATTTCTCGCGTGCCTGTGATGATGACAGCTTTTGTGAAAAATTT 1246
 QY 220 euProTrpLeuAspLysSerValAspPheIleGluYrValAlaYrAspAnGlyCysV 240
 Db 1247 TGCCTGGTGTGACAAATCATGATTTCTATGAGAAAGCAAAAGCTCCAAATGATGATG 1306
 QY 240 alLeuValHsCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaYrIleu 260
 Db 1307 TTTAGTGTGACATGTTTATGCTGGGATCTCCCGCTCCGCAACATCGCTATCGCTATCATCA 1366
 QY 260 euLysArgMetAspMetSerLeuAspGluAlaYrArgPheValLysGluYrValArgPro 280
 Db 1367 TGAAGAGATGACATGCTTTAAGATGAAGCTTACGATTGTGAAAAGAAAAAGCTCA 1426
 QY 280 hrIleSerProAnPheAnPheLeuGlyGlnLeuLeuAspYrGluYrLysIleLys 300
 Db 1427 CTTATCTCCAACTTCAATTTCTTGCGCAACTCTGCACTGTGAGAAAGAAATTAACA 1486
 QY 300 anGlnThrGlyIleSerGlyProLysSerLysLeuLysLeuAnIleLeuGluYrPro 320
 Db 1487 ACCAGACTGGAGATCAGGGCCAAAGCAAACTCAAGCTGGCTGCACTGGAGAGCCAA 1546
 QY 320 anGluProValProAlaValSerGluGlyGlyLysSerGluThrProLeuSerPro 340
 Db 1547 ATGAACTTCTCTGCTGTCTCAGAGGTGACAGAAAAGCGAGAGCCCTCAGTCCAC 1606
 QY 340 roCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHsProAlaSerV 360
 Db 1607 CTTGTCCCACTCTGCTTACTCAGAGCGACAGAGCAAAAGCCCGGTGATCCCGCCAGG 1666
 QY 360 alProSerValProSerValGlnProSerLeuLeuGluAnAspSerProLeuValGlnAla 380
 Db 1667 TGCCTGCGTGGCCAGCGTGCAGCCGCTGCTGTATAGAGACAGCCCGCTGTATCAGGCC 1726

QY 380 euSerGlyLeuHsIleuSerAlaAspArgLeuGluAnAspSerAnLysLeuLysArgSerP 400
 Db 1727 TCAGTGGGTGACCTGTCCGACAGGCTGAAAGACGAAATAGCTCAAGCGTCTCT 1786
 QY 400 heSerLeuAspIleLysSerValSerYrSerAlaSerMetAlaIleSerLeuHsIleY 420
 Db 1787 TCTCTGTGATATCAATACAGTTTCAATATTCAGCCAGCATGGCAGCATCTTATCAGGCT 1846
 QY 420 heSerSerSerGluAnpAlaLeuGluYrYrLysProSerThrThrLeuAspGlyThra 440
 Db 1847 TCTCTCATCAGAAAGATGCTTTGAAATACTACAAACCTTCCACTCTGTGATGGAGCA 1906
 QY 440 anLysLeuGluGlnHsSerProValGlnGluLeuSerGluGluThrProGluThrSerP 460
 Db 1907 ACAAGCTATGACAGATCTCTCCCTGTTCAGAACTATGGAGCAAGATCCCGAAACCAAGTC 1966
 QY 460 roAspLysGluGluIleAspIleProLysLysLeuGlnThrAlaArgProSerAspSerG 480
 Db 1967 CTGATTAAGAGAGAACCCAGCATCCCAAGAGCTCAGACCCCGAGCTTACAGACGCC 2026
 QY 480 IseLysArgLeuHsIseValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 Db 2027 AGAGCAAGCGATTCATTCGTGATCAGAACACAGACAGATGGCACCCCGAGAGTCCCTT 2086
 QY 500 euSerProLeuHsIArgSerGlySerValGluAnAspAnThrIshThrSerPheLeuPheG 520
 Db 2087 TATCTTCATGTGATGAAAGTGGAGCGTGGAGAGCAATTAACACACACAGCTTCTTTTTCG 2146
 QY 520 IYLeuSerThrSerGlnGlnHsIleuThrLysSerAlaGlyLeuGlyLeuLysGlyTrpH 540
 Db 2147 GCCTTCCACCGACCGACGACGACCTCAGAGATCTGCTGGCTGGGCTTAAAGGCTGGC 2206
 QY 540 IseSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpYrPhe 560
 Db 2207 ACTCGAATATCTTGGCCCCCGACAGCTCTAACCCCTTCCCTGACAGAGCTGTATTTTG 2266
 QY 560 IatThrGluSerSerHsPheYrSerAlaSerAlaIleYrGlyGlySerAlaSerYrS 580
 Db 2267 CCACAGAGCTTCACACATCTTACTGCTGCTGACCATCAACGAGGAGAGTGCAGTACT 2326
 QY 580 eAlaYrSerCysSerGlnLeuProThrCysGlyAspGlnValYrSerValArgArg 600
 Db 2327 CTGCTTACGCTGCGACCGACGCTGCCACTTGGGAGAACCAAGTCTATTTCTGCGCAGGC 2386
 QY 600 IrgLysProSerAspArgAlaAspSerArgArgSerTrpHsIleGluGluSerProPheG 620
 Db 2387 GGCAGAGCGCAAGTACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGACCCCTTTG 2446
 QY 620 IuLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGlu 640
 Db 2447 AAAAGCAGTTTAAAGCGAGAGCTGCCAAATGGAATTTGGAGAGCATCATGTCAGAGA 2506
 QY 640 anArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetG 660
 Db 2507 ACAAGTCAAGGAGAAAGCTGGGGAAAGTGGGCACTGACTTACCTTTTGGGACAGATGG 2566
 QY 660 IuIleIleGluValSer 665
 Db 2567 AAATCATTAGAGTCTCC 2583
 RESULT 21
 ABV26826
 ID ABV26826 standard; cDNA, 5145 BP.
 XX
 AC ABV26826;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 26817.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.
OS WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JB;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1, Page 5420-5421, 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (1) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (1) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
XX
XX Alignment Scores:
XX Pred. No.: 0 Length: 5145
XX Score: 472.00 Matches: 663
XX Percent Similarity: 99.40% Conservative: 0
XX Best Local Similarity: 99.40% Mismatches: 2
XX Query Match: 70.98% Indels: 4
XX DB: 5 Gaps: 0
XX
XX US-10-029-345a-109 (1-665) x ABV26826 (1-5145)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGln 20
DB 589 ATGGCCCATGAGTATGATGGAATCAATTTACTGAGAGGTGTGTGCTGTGCTGGAA 648
QY 21 SerGlyThrGlnIleValLeuLeuIleAspSerArgProPheValGlnIlyrAsnThrSer 40
DB 649 AGTGGAAACGGAAAAAGTGTCTGCTAATTGATAGCCGGCCATTGTGGAAATACATAATCATCC 708
QY 41 HisIleLeuGlnAlaIleAsnIleAsnGlySerIleLeuMetIleArgLeuGlnGln 60
DB 709 CACATTTTGGAAACCTTATATATCACTGCTCCAGCTTATGAGGAAAGTTTGCAACG 768
QY 61 AspIleValLeuIleThrGlnLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
DB 769 GACAAAGTGTATTTACAGAGCTCATCCGACATTGACGCAACATAGAGTTGACATTGAT 828
QY 81 CysSerGlnIleValValValIlyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 829 TGCAGTCAGAAAGGTGTAGTTTACATCAAAAGCTCCCAAGAGTGTCTCTCTCTTCA 888

QY 101 AspCysPheLeuThrValIleLeuGlnIlyrLeuGlnIlyrSerPheAsnSerValHisIleu 120
DB 889 GACTGTTTCTCAGCTGATCTTCTGGGTAACTGGAGAAAGACTTCAACTCTGTTCACTCG 948
QY 121 LeuAlaGlyIlyrPheAlaGluPheSerArgCysPheProGlyLeuCysGlnIlyrIlyrSer 140
DB 949 CTTCAGAGTGGGTGTGCTGAGATTCCTGTTGTTCCTGAGCTCTCTGAGAGAAATCC 1008
QY 141 ThrIleValProThrCysIleSerGlnProCysIleuProValAlaAsnIleGlyProThr 160
DB 1009 ACTCTAGTCCCTACTGATTCATTTCTCAGCTGTGCTTACTGTTGCAACATTTGGCCCAACC 1068
QY 161 ArgIleLeuProAsnLeuTyrIleuGlyCysGlnArgAspValLeuAsnIleGlnLeuIle 180
DB 1069 CGAATTCCTCCCAATCTTATCTTGCTGCCAGGAGATGCTCTCAACAGAGACTGAT- 1127
QY 181 -GlnGlnAsnGlyIlyrIlyrValIleuAsnAlaSerTyr-ThrCysProIlyrProAsp 200
DB 1128 GCAGCAGATGGAGATTGGTTATGTGTAAATGCCAGCAA-TACCTGTCCAAAGCCTGACT 1186
QY 200 helIeProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlnIlyrIle 220
DB 1187 TTATCCCGAGTCTATTTCTCGGTGTGCTGTGAATGACAGCTTTGTGAGAAATTT 1246
QY 220 euProTyrLeuAspIlyrSerValAspPheIleGlyIlyrAlaIlyrAlaSerAsnGlyCys 240
DB 1247 TCCCGTGTGTGGACAAATCAGTATGATTTCAATTGAGAAAGCAAAAGCCTCCATGATGATG 1306
QY 240 AlIleValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleu 260
DB 1307 TTCTAGTCCAGCTGTTTACTGAGATCTCCGCTCCGACACATCTCTATCGCTACATCA 1366
QY 260 eIlyrArgMetAspMetSerLeuAspGlnAlaTyrArgPheValIlyrGlnIlyrArgProT 280
DB 1367 TGAAGAGATGGACATGCTTTAGTAAAGCTTACAGATTTGTGAAAGAAAAAGACCTA 1426
QY 280 hIleSerProAsnPheAsnPheLeuGlnIleLeuAspTyrGlnIlyrIlyrIlyrIlyr 300
DB 1427 CTATATCTCCAAACTCAATTTTCTGGGCCAACTCTGAGCTATGAGAAAGATTAGA 1486
QY 300 SngIlnThrGlyAlaSerGlyProIlyrSerIlyrIleuIlyrLeuIleHisIleuGlnIlyrPro 320
DB 1487 ACCAGACTGGAGCATCAGGCGCAAGCAAACTCAGCTGTGACCTGGAGAGCCAA 1546
QY 320 SngIlnProValProAlaValSerGlnIlyrIlyrIlyrSerGlnIlyrProLeuSerProp 340
DB 1547 ATGAACCTGTCCTCTCTCTCTCAGAGGTGAGACAGAAAGCAAGCGCCCTCAGTCCAC 1606
QY 340 roCysAlaAspSerAlaThrSerGlnAlaIleGlyGlnArgProValHisProAlaSerV 360
DB 1607 CCTGTCCGACTCTGCTACCTCAGAGCAGCAGCAAGAGCCCTGTGATCCCGCCAGCG 1666
QY 360 AlProSerValProSerValGlnProSerIleuGlnAspSerProIleuValGlnAla 380
DB 1667 TGCCAGCGGTGCCAGCGGTGCGCTGTGAGAGACGCCGCTGGTACAGGGCGC 1726
QY 380 euSerGlyLeuHisIleuSerAlaAspArgLeuGlnAspSerAsnIlyrIlyrIlyrIlyr 400
DB 1727 TCAGTGGCTGACCTGTCCGACAGCAGCTGGAAGACAGCAATAGCTCAAGCTTCT 1786
QY 400 heSerLeuAspIleIlyrSerValSerTyrSerAlaSerMetAlaIleSerLeuHisIlyr 420
DB 1787 TCTCTCTGATATCAATCAAGTTTATATTGAGCAGCATGGCAGCATCTTACATGAGCT 1846
QY 420 heSerSerSerGlnAspAlaLeuGlnIlyrTyrIlyrProSerThrThrIleuAspGlyTyr 440
DB 1847 TCTCTCTCAACAAATCTTGTGAATCTCAAACTTCCACTGATGAGTGGACCA 1906
QY 440 snIlyrLeuCysGlnPheSerProValGlnIlyrLeuSerGlnIlnThrProGlnThrSerP 460
DB 1907 ACAAGCTATGCCAGATTCCTCTCTCTCAGGAACATTCGAGAGAGACTCCCAAAACAAGTC 1966
QY 460 roAspIlyrGlnIlyrAlaSerIleProIlyrIlyrLeuGlnThrAlaArgProSerAspSerg 480

```

Db      1967 CTGATAGAGGAGGAGGAGGATCCCAAGAGCTGAGACCGCCAGCCTTCAGACAGC 2026
Qy      480 InSerIyArGLeuHISerValArgThSerSerSerGlyThraIaGInaYSerLeuL 500
Db      2027 AGGCAAGGAGGATTCATTCGCTCAGAACGACGAGCTGACCCGCCAGAGGATCCCTTT 2086
Qy      500 euseProLeuHISaYSerGlySerValGluuApAnYrThsThSerPheLeuPhg 520
Db      2087 TATCTCAGCTGCAATCGAAGGGAGCGGTGAGGACAAATTAACACACAGCTTCCTTTTG 2146
Qy      520 IyLeuSerThSerGInGInHISerLeuThLySerSerIaGlyLeuGlyLeuYrPh 540
Db      2147 GCCTTCCACACGACGACGACACCTCACAAAGCTCTGCGCCCTTAAGGCTGCG 2206
Qy      540 IASerAspIleLeuAlaProGInThSerThProSerLeuThSerSerTyrPheA 560
Db      2207 ACTCGATATCTTGCGCCCGCCAGACCTTAACCCCTTCCTGACACGACGCTGATTTTG 2266
Qy      560 IAThGluSerSerHISpHeTySerSerAlaSerAlaIleTyrglyGlySerAlaSerTyrS 580
Db      2267 CCACAGATCTTACACTTCTACTCTGCTCAGCCCATCTACGAGGAGGAGGAGGAGTACT 2326
Qy      580 eAlaTySerCySerSerGInLeuProThCySerIyAspGInValTySerValArgArg 600
Db      2327 CTGCCACAGCTGACGACGACGCTGCCCATCTTGCGGAGACCAAGCTTATCTGCGCAGC 2386
Qy      600 rGInLyArProSerArPaArgAlaAspSerArgYrSerTyrHISGInGInSerProPhg 620
Db      2387 GCGAGAGGCAAGTGCACAGAGCTGACTCGCGGCGGAGCTGGCATGAAGAGAGCCCTTTG 2446
Qy      620 IuLySerIuPheLyArGaYrSerCySerGInMetGInuHeglyGlyuSerIleMetSerGlu 640
Db      2447 AAAAGAGTTTAAACGACAGAGCTGCCAATGTGAATTTGAGAGAGACCATCATGTGAGAA 2506
Qy      640 sArGSerArgGInGInLeuGlyLyValAlGlySerGInSerSerPheSerGlySerMetG 660
Db      2507 ACAGGTCACGCGAAGAGCTGCGGAAAGTGGGAGTACGTACGCTTTTGGGCGACGATG 2566
Qy      660 IuIleIleGluValSer 665
Db      2567 AAATCATGTAGGCTCC 2583

RESULT 22
ABV27131
ID      ABV27131 standard; cDNA, 5145 BP.
XX
XX      ABV27131,
AC      16-SEP-2002 (first entry)
XX
DT      Human prostate expression marker cDNA 27122.
XX
DE      Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX      Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX      pharmacogenomic marker; gene; ss.
XX      Homo sapiens.
XX      OS
XX      WO200160860-A2.
XX      PN
XX      23-AUG-2001.
XX      PD
XX      20-FEB-2001; 2001WO-US005171.
XX      PF
XX      17-FEB-2000; 2000US-0183119P.
XX      PR      16-MAR-2000; 2000US-0189862P.
XX      PR      25-MAY-2000; 2000US-0207454P.
XX      PR      09-JUN-2000; 2000US-0211314P.
XX      PR      18-JUL-2000; 2000US-0219007P.
XX      PR      13-DEC-2000; 2000US-0255281P.
XX
XX      (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX

```

```

PI      Schlegel R, Endege WO, Monahan JB;
XX
XX      WPI; 2001-662795/76.
XX
PT      Novel isolated nucleic acid molecule associated with cancerous state of
XX      prostate cells and correlating with presence of prostate cancer, useful
XX      for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS      Claim 1, Page 5499-5500; 11750pp; English.
XX
XX      The invention relates to an isolated nucleic acid molecule (1) comprising
XX      a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
XX      specification or its complement. (1) is useful for: (a) assessing whether
XX      a patient is afflicted with prostate cancer; (b) monitoring the efficacy
XX      of a test compound to inhibit prostate cancer in a patient; (c) assessing the
XX      efficacy of a therapy for inhibiting prostate cancer in a patient;
XX      (d) selecting a composition for inhibiting prostate cancer in a patient;
XX      (e) assessing the prostate cell carcinogenic potential of a compound; (f)
XX      determining whether prostate cancer has metastasized in a patient; (g)
XX      assessing the aggressiveness or indolence of prostate cancer in a patient
XX      ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ      Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
XX
XX      Alignment Scores:
XX      Score: 0 Length: 5145
XX      Percent Similarity: 472.00 Matches: 663
XX      Best Local Similarity: 99.40% Conservative: 0
XX      Query Match: 70.98% Mismatches: 2
XX      DB: 5 Gaps: 4
XX
US-10-029-345A-109 (1-665) x ABV27131 (1-5145)
Qy      1 MetAlaHISerMetIleGlyThrGInIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db      589 ATGGCCCATGAGATGATTTGAACTCAATTTGTTACTGAGAGGTTGCTGCTCGGAA 648
Qy      21 SerGlyThrGluValLeuLeuLeuIleAspSerArgProPheValGluTyrAnThrSer 40
Db      649 AGTGGAGCGGAAAGAGTGTCTTAATGTATGATCCGGCCATTTGTGGATATACATATACCC 708
Qy      41 HisIleLeuGluAlaIleAsnIleAsnCySerIyLeuMetIyArGArgLeuGInGIn 60
Db      709 CACATTTTGGAGGCCATTAATATCACTGCTCAAGCTTATGAAGGAGGTTGCAACAG 768
Qy      61 AspLyValLeuIleThGluLeuIleGInHISerAlaIleHISerValAspIleAsp 80
Db      769 GACAAAGGTATTAATACAGAGCTCATCCAGATTCAGCGAACAATRAAGGTGACATTGAT 828
Qy      81 CySerGInLyValValValValTyrAspGInSerSerGInAspValAlaSerLeuSer 100
Db      829 TCACGTGACGAAGGTTGTGATTACATCAAACTCCCAAGAGTGGCCCTCTCTCTTCA 888
Qy      101 AspCySpHeLeuThValLeuLeuGlyLyLeuGluLySerPheAsnSerValHISLeu 120
Db      889 GACTGTTTCTCAGCTACTCTCTGGGTAAACTGAGAGAGGCTTCAACTCTGTCACCTG 948
Qy      121 LeuAlaGlyGlyPheAlaGluPheSerArgCyPheProGlyLyLeuCySerGluGlyIySer 140
Db      949 CTGCAAGGAGGAGTGTGAGTCTCTCTGTTTCTGCTGCTGCTGTAAGGAAATATCC 1008
Qy      141 ThrLeuValProThCySerIleSerGInProCyLeuProValAlaAsnIleGlyProTh 160
Db      1009 ACTCTAGTCCCTACCTGATTTCTCAGGCTTGCTTAACGTTGCCAACAATTTGGGCCAAC 1068
Qy      161 ArgIleLeuProAsnLeuTyrLeuGlyCySerGInArgAspValLeuAsnLySerLeu 180
Db      1069 CGAATTCCTCCCAATCTTATCTTGTGCTGCGCCAGAGAGTCTCTCAACAAGAGAGCTGAT- 1127
Qy      181 -GInGInAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCySpPolyProAsp 200

```


CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or incidence of prostate cancer in a patient
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	0	Length:	5145
Score:	472.00	Matches:	663
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	70.98%	Indels:	4
DB:	5	Gaps:	0

US-10-029-345A-109 (1-665) x ABV26923 (1-5145)

```
QY 1 MetAlaHieGluMetIleGlyThrGlnIleValThrGluArgPheValAlaLeuLeuGlu 20
DB 589 ATGGCCCATGAGTATGATTGGAACTCAAAATTGTTACTGAGAGGTGGTGGCTGCTGGAA 648
QY 21 SerGlyThrGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 649 AGTGAACGGAAAAAGTGGTGAATTGATAGCCGGCCATTGTGGAAATACATATACATCC 708
QY 41 HisIleLeuGluAlaIleAsnIleAsnCySerIleLeuMetIleArgIleGlnGln 60
DB 709 CACATTTTGGAAACCATTAATATCAACTGCTCCAACTTAATGAAGGAAGTTGCAACG 768
QY 61 AspIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
DB 769 GACAAAGTGTATTAACAGACTCATCCAGACTTCAGCAACATTAAGGTTGACATTGAT 828
QY 81 CySerGlnIleValIleValIleValIleValIleValIleValIleValIleValIle 100
DB 829 TGCAGTCAGAGGTGATGATTACGATCAAAAGCTCCCAAGATGTGCTCTCTCTTCA 888
QY 101 AspCyPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
DB 889 GACTGTTTCTCACTGACTTCTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACTG 948
QY 121 LeuAlaGlyIlePheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 140
DB 949 CTTCGACGGTGGTTCGAGTCTCTCTGTTGTTTCCCTGGCTCTGTGAAGAAATCC 1008
QY 141 ThrLeuValProThrCyAlaSerGlnProCyAlaProValAlaAsnIleGlyProThr 160
DB 1009 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTGCTGCTGCAAACTGGGCCAAC 1068
QY 161 ArgIleLeuProAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
DB 1069 CGAAATCTTCCCAATCTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1127
QY 181 -GlnGlnAsnGlyIleGlyThrValIleAsnAlaSerTyR-ThrCyPheProValAsp 200
DB 1128 GCGAGCAAGATGGATGTTATGTTAAATGCCAGAA-TACCTGTCGCAAACTGAGCT 1186
QY 200 HeIleProGluSerHisPheLeuArgValProValAsnAspSerPheCyGluIleIle 220
DB 1187 TTATCCCGGAGTCTCATTTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246
QY 220 eubProTrpLeuAspIleSerValAspPheIleGluIleValAlaIleAsnIleGly 240
DB 1247 TGCCTGGTGGTGGCAATCATGATTTCAATTGAGAAAGCAAAAGCTCCAAATGGATGTC 1306
QY 240 AlLeuValHisCyLeuAlaGlyIleSerArgSerIleThrIleAlaIleAlaTyR-Ile 260
DB 1307 TTCTAGTGCATTTAGCTGGGATCTCCCGCTCCGACACATGCGCTATGCTTACATCA 1366
QY 260 eLysArgMetAspMetSerLeuAspGluAlaTyRArgPheValIleGluIleArgPro 280
DB 1367 TGAAGGAGATGAGCATGCTTTTATGATGAGCTTACAGATTTGTGAAGAAAAAGACCTA 1426
```

```
QY 280 HisIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyRGlulysIleIleVal 300
DB 1427 CTATATCTTCAAACTTCAATTTTCTGGCCAACTCTGAGCTATGAGAAAGATTAAGA 1486
QY 300 eGlnThrGlyAlaSerGlyProIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 320
DB 1487 ACCAGCTGAGCATCAGGCGCCAAAGAGCAAACTCAAGCTGCTGCACTGGAGAACCA 1546
QY 320 eGlnProValProAlaValSerGluGlyGlyGlnIleValSerGluThrProLeuSerPro 340
DB 1547 ATGAACCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1606
QY 340 roCyAlaAspSerIleThrSerGluAlaGlyIleGluArgProValHisProAlaSerTy 360
DB 1607 CTTGTGCGACTCTGCTTACTCAGAGGCGAGGAGCAAGAGGAGGAGGAGGAGGAGGAGG 1666
QY 360 AlProSerValProSerValGlnProSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 380
DB 1667 TGCACAGGTGCGCCAGCGTGCAGCGCTGCTGTTAGAGGACAGCCGCTGTACAGGCC 1726
QY 380 eSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnIleValIleValIle 400
DB 1727 TCAGTGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1786
QY 400 HeSerLeuAspIleIleSerValSerTySerIleSerIleAlaIleSerLeuHisGlyP 420
DB 1787 TCTCTGATATCAATCATGATTTCAATTCACGACGATGAGCATCTTATCAATGAGCT 1846
QY 420 HeSerSerSerGlnAspAlaLeuGluTyRThrIleValSerThrThrLeuAspGlyThra 440
DB 1847 TCTCTCATCAGAAAGTGTGTTGGAATATCAAACTTCCACTGATGAGTGGAGCA 1906
QY 440 eNlyLeuCyGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerP 460
DB 1907 ACAAGCTAGCAGATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1966
QY 460 roAspIleGluGluIleSerIleProIleValLeuGlnThrIleArgProSerAspSerG 480
DB 1967 CTGATTAAGAGGAGGAGCAGCATCCCAAGAGCTGCAAGCGCAGAGCTTTCAGAGCC 2026
QY 480 InSerIleArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
DB 2027 AAGAGCAAGGATTTGATTCGTCAGAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 2086
QY 500 eSerProLeuHisArgSerGlySerValGluAspAsnTyRHisThrPheLeuPheG 520
DB 2087 TATCTTCATGATGAGATGAGGAGCTGAGAGCAATTAACCAACCACTTCTCTTTTGC 2146
QY 520 ILeuSerThrSerGlnGlnIleLeuThrIleValSerIleAlaIleLeuIleValIle 540
DB 2147 GCTTTTCCACGAGCCAGCAGCACTTACGAACTGCTGCTGCTGCTGCTGCTGCTGCT 2206
QY 540 ILeuSerIleLeuAlaProGlnThrSerThrProSerLeuThrPheSerIleTyR-Phea 560
DB 2207 ACTCGATATCTTGGCCCCCAGACCTTACCTTCCCTGACAGCAGCAGCTGATTTTGG 2266
QY 560 IeThrGluSerSerHisPheTyRSerAlaSerAlaIleTyRGlulysSerAlaSerTyR 580
DB 2267 CCACAGAGTCTCACACTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2326
QY 580 eRAlaTyRSerCySerGlnLeuProThrCyGlyIleAspIleValIleIleValIleArg 600
DB 2327 CTGCTTACAGCTGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2386
QY 600 rGlnIleProSerAspArgAlaAspSerArgSerTrpHisGluGluSerProPheG 620
DB 2387 GCGAGAGCCAGTACAGAGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2446
QY 620 IuIleGlnPheIleArgArgSerCyGlnMetGluPheGlyGluSerIleMetSerGluAl 640
DB 2447 AAAAGCAGTTTAAACGAGAGCTGCAATGGAATTTTGAAGAGGACATGTCAGAGA 2506
QY 640 eNArgSerArgIleGluIleGlyIleValIleIleSerGlnSerPheSerGlySerMetC 660
```


|||||
Db 2507 AACGGTCACGGGAGAGCTGGGAAAGTCGGGACGCTAGCTTTTGGGAGAGATGG 2566
QY 660 lulllelllegluValSer 665
Db 2567 AATCATTTGAGGCTTCC 2583
RESULT 24
ABV27135
ID ABV27135 standard; cDNA; 5145 BP.
XX
AC ABV27135;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 27126.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-018319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 5502-5503; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
XX
Alignment Scores:
Pred. No.: 0 Length: 5145
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 5 Gaps: 0
US-10-029-345A-109 (1-665) x ABV27135 (1-5145)
QY 1 MetaAnalysisMetileGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20

|||||
Db 589 ATGGCCCATGAGATGATTGGAACTCAATTGTGATGTAGAGAGTTGGTCTGCTGCA 648
QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db 649 AGTGGAAAGCAAAAAGTCTGCTGATGATGATGACCGGCGCATTTGTGAATACATCAATCC 708
QY 41 HistIleLeuGluAlaIleAsnIleAsnGlySerIleuMetIleArgArgLeuGlnGln 60
Db 709 CACATTTTGGAAAGCCATTATATCAATGCTCCAGCTTCAAGATTGTAAGAGGTTGCAACAG 768
QY 61 AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleValAspIleAsp 80
Db 769 GACAAAGCTTATATACAGAGCTATCCAGCATTCAGGAAACATTAAGTTGACATTGAT 828
QY 81 CysSerGlnIleValValIleTyrAspGlnSerSerGlnAspValAlaSerIleuSerSer 100
Db 829 TGCAGTCAGAAAGTTGTGATTACGATCAAGCTCCAAAGATGTTGCCCTCTCTCTTCA 888
QY 101 AspCysPheLeuThrValLeuLeuGluIleuGluIleuSerPheAsnSerValHisIleu 120
Db 889 GACTGTTTCTCAGCTGACTCTTGAGTAACTGGAGAAAGACCTTCACTGTTCACTG 948
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCysGluGlyLysSer 140
Db 949 CTTCAGAGTGGGTTTGTGAGTTCTGCTGTTTCCCTGGGCTCTGTGAAGAAATCC 1008
QY 141 ThrIleuValProThrCysIleSerGlnProCysIleuProValAlaAsnIleGlyProThr 160
Db 1009 ACTCAGTCCCTACCTGATTCATTCAGCTTACCTGTTACCTGTGGCCAAATGGGCAACC 1068
QY 161 ArgIleuProAsnLeuTyrLeuGluIleuGlnIleuArgPheValLeuAsnIleGluIle 180
Db 1069 GCAATCTTCCCAATCTTATCTTGCTGCGCAGGAGATGTCCTCAACAGAGAGCTGAT- 1127
QY 181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProLysProAsp 200
Db 1128 GCAGCAGAAATGGATTTGTTATGTTAAATGCCAGCAA-TACCTGTCCAAAGCTGACT 1186
QY 200 helIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle 220
Db 1187 TTATCCCGGAGATCTCATTTCTCGCGTGGCTCGGATGACAGCTTTGTGAGAAATTT 1246
QY 220 eupProThrLeuAspLysSerValAspPheIleGluValAlaIleSerAsnGlyCysV 240
Db 1247 TGCCTGTGTGACAAATCAAGTACATTCATTGAGAAAGCAAAAGCTTCAATGAGATGTG 1306
QY 240 AlIleuValHisCysIleuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaIleAla 260
Db 1307 TTCTAGTGCACATGTTTAGCTGGATCTCCGCTCCGACACATGCTATGCTTACATCA 1366
QY 260 eLysArgMetAspMetSerIleuAspGluAlaIleTyrArgPheValIleGluLysArgPro 280
Db 1367 TGAAGAGATGAGACATGCTTTAGATGAGCTTACAGATTTGTGAAAGAAAAAGACCTTA 1426
QY 280 hrIleSerProAsnPheAsnPheLeuGluIleuLeuAspTyrGluLysIleValIleVal 300
Db 1427 CTATATCTCCAAACTCAATTTCTGGGCAACTCTGGACTATGAGAAAGATTAAGA 1486
QY 300 snGlnThrGlyAlaSerGlyProLysSerIleuLysIleuLeuHisIleuGluLysPro 320
Db 1487 ACCAGACTGAGGATTCAGGGGCCAAAGAGCAAACTCAAGCTCTCTGACCTGGAGAAAGCAA 1546
QY 320 snGluProValProAlaValSerGluGlyIleuLysSerGluThrProLeuSerPro 340
Db 1547 ATGAACCTGTCCTCGCTGCTCAAGAGGATGAGAAAGAGCAAGCCCTCTGCTGCTCAC 1606
QY 340 roCysAlaAspSerAlaThrSerGluAlaIleGluIleuArgProValHisProAlaSer 360
Db 1607 CTTGTGCGGACTCTGCTACCTCAGAGGAGAGAGCAAAAGCCCTGTGATCCCTCCAGCG 1666
QY 360 AlProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAla 380


```
Db      1667 TGCCACGCGTCCAGCGTCAGCGCTGCTTTAGAGGACAGCCCGCTGTACAGAGCCG 1726
Qy      380 eUSerGlyLeuHISLeuSerAlaAspArgLeuGluAspSerAnlyISLeuValArgSerP 400
Db      1727 TCAGTGGGCTGCACCTGTCGAGACAGAGCGTGGAAACAGACATTAAGCTTCAGCGTTCTCT 1786
Qy      400 hSeSerLeuAspIleuValSerValSerTyrSerAlaSerMetAlaAlaSerLeuHISGlyP 420
Db      1787 TCTCTTGGATATCAATCATAGTTTCTATTCAGCCAGCATGGACATCTTTATCATAGGCT 1846
Qy      420 hSeSerSerSerGluAspAlaLeuGluTyrTyrIlyAspProSerThrThrLeuAspGlyThra 440
Db      1847 TCTCCATCAGAAAGATGCTTTGGAAATCATCAAACTTCCTCACTACTCGATGGAGACCA 1906
Qy      440 emlyLeuCySerGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerP 460
Db      1907 ACAAGCTATGCCAGTCTCTCCCTGTTCAGGAACTATCGGAGCGAGATCCCGAAACAGTC 1966
Qy      460 roAspIlySerGluAlaSerIleProIlyLeuGluGlnThrAlaArgProSerAspSerG 480
Db      1967 CTGATTAAGAGAGAGCCAGCATCCCAAGAGCTGGACAGCCGACAGCTTCAGACAGCC 2026
Qy      480 InSerIlyAspArgLeuHISerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuL 500
Db      2027 AGAGCAAGGAGATTGCATTGGGTGAGAACAGACAGAGCGAGCGGACCGCCAGAGGTCCTTT 2086
Qy      500 eUSerProLeuHISArgSerGlySerValGluAspAnlyThrISThrSerPheLeuPheG 520
Db      2087 TATCTCCACTGCATCCAGAGGAGCGGTGAGAGACATTAACACACAGCTTCCTTTTCG 2146
Qy      520 IlyLeuSerThrSerGlnGlnHISLeuThrIlySerValISGlyLeuGlyLeuValGlyThP 540
Db      2147 GCTTTTCCACCGACCGACGACCTTCAAGAGTCTGCTGGCTGGCTTAAAGGCTGGC 2206
Qy      540 IASerAspIleLeuAlaProGlnInThrSerThrProSerLeuThrSerSerTyrIlyPhea 560
Db      2207 ACTCGATATCTTGGCCCCCGACCACTTACCCCTTCCTCGACAGACAGCTGTGATTTTG 2266
Qy      560 IATHrGluSerSerHISpHeTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrS 580
Db      2267 CCACAGAGTCTTCACTTACTCTGCTGCTCAGCCATCTACGAGAGCAGTGCACAGTTACT 2326
Qy      580 eAlaIlySerCySerSerGlnLeuProThrCySerGlyAspGlnValTyrSerValArgArg 600
Db      2327 CTGCTTACAGCTGACAGCGACCTGCCACTTGGGAAACAGTCTATCTGTGCGAGG 2386
Qy      600 rGlnIlyAspProSerAspArgAlaAspSerArgArgSerTyrPheGlnGluSerProPheG 620
Db      2387 GCGAGAAAGCCAACTGACAGAGCTGATCGCGGCGAGCTGGCATGAAGAGAGCCCTTTG 2446
Qy      620 IlyLeuGlnPheIlyAspArgSerCySerGlnMetGluPheGlyGluSerIleMetSerGlu 640
Db      2447 AAAAGCAGTTTAAACGACAGAGCTGCCAAATGGAAATTTGGAGAGACATCATGTGAGA 2506
Qy      640 sAspArgSerArgGluGluLeuGlyIlyValGlySerGlnSerSerPheSerGlySerMetG 660
Db      2507 ACAGGTCACGGGAAAGCTGGGGAAAGTGGGAGTCACTTACCTTTTGGGAGAGATGG 2566
Qy      660 IuIleIleGluValSer 665
Db      2567 AAATCATTAAGTCTCC 2583
```

```
KM      pharmacogenomic marker; gene; ss.
XX
OS      Homo sapiens.
XX
PN      WO200160860-A2.
XX
PD      23-AUG-2001.
XX
PF      20-FEB-2001; 2001WO-US005171.
XX
PR      17-FEB-2000; 2000US-0183119P.
PR      16-MAR-2000; 2000US-0189862P.
PR      25-MAY-2000; 2000US-0207454P.
PR      09-JUN-2000; 2000US-0211314P.
PR      18-JUL-2000; 2000US-0219007P.
PR      13-DEC-2000; 2000US-0255281P.
XX
PA      (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI      Schlegel R, Endege WO, Monahan JB;
DR      WPI; 2001-662795/76.
XX
PT      Novel isolated nucleic acid molecule associated with cancerous strate of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS      Claim 1; Page 6009-6010; 11750pp; English.
XX
CC      The invention relates to an isolated nucleic acid molecule (1) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (1) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the efficacy of progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ      Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
XX
Alignment Scores:
Pred. No.: 0 Length: 5145
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 5 Gaps: 0
XX
US-10-029-345A-109 (1-665) x ABV28657 (1-5145)
Qy      1 MetAlaHISGluMetIleLeuIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db      589 ATGGCCCAAGAAATGATTTGAACTCAAAATTTGATGAGAGGTGGTGGCTGCGTGGAA 648
Qy      21 SerGlyThrGluValLeuLeuIleAspSerArgPProPheValGluTyrAnthrSer 40
Db      649 AGTGAACGGAAAAAGCTGCTGTAATGTATACCGGCCCATTTTGTGAATACATATCC 708
Qy      41 HISLeuGluAlaIleAnlyLeuAnlyLeuAnlyLeuAnlyLeuAnlyLeuAnlyLeuAnly 60
Db      709 CACATTTTGAAGCCATTAATATCACTCTCCAAAGCTTATGAGAGCGAAGTTGCAACAG 768
Qy      61 AspIlyValLeuIlePheGluLeuIleGlnHISerAlaIlyValIleValIleValIleAsp 80
Db      769 GACAAAGTGTATTAACAGAGCTCATCCAGACTTACGCAAAACATTAAGGTGACATTAAT 828
Qy      81 CySerGlnIlyValValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db      829 TGCAGTCAGAAAGTGTATTAATTTTACATCAAAAGCTCCCAAGAGTTCCTCTCTCTTCA 888
```

QY	101	AspGpPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu	120
Db	889	GACTGTTTTCTCACTGACTTCTTGGGTAAACGGAAGAGCTTCAACTCTGTCACTG	948
QY	121	LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuCysGlnGlySer	140
Db	949	CTTGCAAGGGGTTGGTCTAGTTCCTCTGTTTTCCTCGGCTCTGGAGAGAAATCC	1008
QY	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr	160
Db	1009	ACTAGTACCTCACTGACATTTCTCAGCCTTCTTACCTGTTGCGCAATTTGGCCAAAC	1068
QY	161	ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeu	180
Db	1069	CAAAATTTCTCCAAATCTTTATCTTGGCTGCCAGCAAGATGTCTCAAAAGAGCTGAT-	1127
QY	181	-GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProLysProAsp	200
Db	1128	GGAGCAGATGGGATTTGGTTATGTGTAAATCCAGCAA-ThrCTGTCCAAAGCTGACT	1186
QY	200	heIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle	220
Db	1187	TTATCCCCAGCTCATTTCTCGCTGGCTGATATGACAGCTTTGTGGAAAAATTT	1246
QY	220	euProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysV	240
Db	1247	TGCGGTGGTGGCAAAATCAGTAGATTTCAATGAAAGAAAGAAACCCTCCATGATGTG	1306
QY	240	AlLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleM	260
Db	1307	TTCTAGTGACAGTTTAGCTGGAGCTCCCGCTCCGCCACCATCGCTATCCCTACATCA	1366
QY	260	eIlyArgMetCaspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProt	280
Db	1367	TGAAGAGATGACATGTCTTATGTAAGCTTACAGATTTGTGAAGAAAAAGAAAGACTTA	1426
QY	280	hrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLysIleLysA	300
Db	1427	CTATATCTCCAAACTTCAATTTTCTGGGCCAAACTCTTGAGCTATGAGAGAAATTAAGA	1486
QY	300	sngInThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGluLysProA	320
Db	1487	ACCAGACTCGAGACATCAGGGCCAAAGCAAACTCAAGCTGTCTGACCTGGAGAGCCAA	1546
QY	320	sngInuProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProp	340
Db	1547	ATGAACCTGTCTCTGTCTCTCAAGGGGTGACAGAAAGGAGACGCCCCCTCAGTCCAC	1606
QY	340	roCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerV	360
Db	1607	CCTGTGCCGACTCTCTCACTCAAGGACAGACAGCAAAAGCCCGCTGATCTCCGCCAGCG	1666
QY	360	AlProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaL	380
Db	1667	TGCCAGAGCTCCAGCGGTGCAGCGCTGTGTTAAGAGACAGCCCGCTGGTACAGGCGC	1726
QY	380	eusSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerP	400
Db	1727	TCAGTGGGCTGACCTGTCTCGCAACAGGCTGGAAGACAGCAATTAACCTCAAGGTTTCT	1786
QY	400	heserLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyP	420
Db	1787	TCTCTCGGATCAATCAATCAGTTTCAATTCAGCCACAGATGGACAGATCTTACATGGCT	1846
QY	420	heserSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThra	440
Db	1847	TCTCTCATATCGAAGATGCTTTGGAAATTAACAACCTTCACTCACTCGATGGAGCCA	1906
QY	440	snuLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerP	460
Db	1907	ACAAAGCTATGCAATTTCTCCCTGTGTGAAGAACTATGAGACAGACTCCCGAAACCAAGTC	1966

[illegible]

XX Schlegel R, Endege WO, Monahan JE.
PI WPI, 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
PS Claim 1, Page 4033-4034, 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;
SQ
Alignment Scores:
Pred. No.: 0 Length: 5145
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 5 Gaps: 0
US-10-029-345A-109 (1-665) X ABV22827 (1-5145)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 589 ATGGCCCAATGAGATGATGAGACTCAAAATTTGTTACTGAGAGGTGGTGGCTGCTGAGAA 648
QY 21 SerGlyThrGluLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 649 AGTGAACCGGAAAAAGTGGCTGCTAATGATAGCCGCGCATTTTGGAAATACATAATCATC 708
QY 41 HisIleLeuGluAlaIleAsnIleAsnCySerLeuLeuMetIleArgArgLeuGlnGln 60
DB 709 CACATTTTGGAGCCCATTAATATCAACTGCTCCAACTTAATGAAAGGAGTTGCAACAG 768
QY 61 AspIleValLeuLeuLeuThrGluLeuLeuIleGlnHisSerAlaIleHisIleValAspIleAsp 80
DB 769 GACAAAGGTGTAATTAACAGAGCTCATCCAGCATTCAGCGAACAATAGGTTGACATTTGAT 828
QY 81 CysSerGlnIleValValValIleTyraAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 829 TCGAGTCAGAGAGTTGATGATTACATCAAGCTCCCAAGATGTTCCCTCTCTCTTCA 888
QY 101 AspCysPheLeuThrValLeuLeuGlyIleGluGlnIleGluSerPheAsnSerValHisLeu 120
DB 889 GACTGTTTCTCACTACTCTCTGGGTAACTGAGAAAGAGCTTCAACTCTGTTCACTG 948
QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuGlyGluGlyIleAspSer 140
DB 949 CTTCGAGGGGGGTTCTGAGTCTCTCGTTGTTTCCCTGGCTCTGTGAAGAAATATCC 1008
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 1009 ACTCTAGTCCCTACCGCATTTCTCAGCCTTGCTTACCTGTTCCAAACATTGGGCCAAC 1068
QY 161 ArgIleLeuProAsnLeuLeuTyriLeuGlyCysGlnArgAspValLeuAsnIleGluLeu 180
DB 1069 CGAATTTCTTCCCAATCTTAATCTTGGCTGCGACGAGATGCTCTCAACAGAGGCTGAT- 1127
QY 181 -GlnGlnAsnGlyIleGlyTyriValLeuAsnIleSerTyriThrCysProLeuProAsp 200

DB 1128 GCAGCAGATGGGATTTGTTATGTTAATCCAGCA-TACTCTCCAAAGCCGACT 1186
QY 200 HeIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGluIleLeu 220
DB 1187 TTATCCCGAGCTCATTTCTCGCGTGCCTGTGAATACAGACTTTTGTGAAATATT 1246
QY 220 euProThrLeuAspIleSerValAspPheIleGluValAlaIleAsnArgIleCys 240
DB 1247 TCCCGGTGGACAAATCAGATTTTATGAGAAAGCAAAAGCTCCAAATGATGTG 1306
QY 240 AlLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyriLeu 260
DB 1307 TTCTAGTGCACCTTTAGTGGAGTCTCCGCTCCGACCACTGATCGCTACATCA 1366
QY 260 eTLeuArgMetAspMetSerLeuAspGluAlaTyriArgPheValIleGlyIleArgPro 280
DB 1367 TGAGAGGATGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAGAACACTA 1426
QY 280 hTIIeSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyriGlyIleValIleVal 300
DB 1427 CTATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGACTATGAGAAAGATTMAA 1486
QY 300 hNIGIThrGlyAlaSerGlyProIleSerLeuLeuLeuLeuHisLeuGluIlePro 320
DB 1487 ACCAGACTGAGACATCAGGCGCAAGAGCAAACTCAAGCTGCTGACCTGGAGAACCA 1546
QY 320 hNIGIProValProAlaValSerGluGlyIleGlnIleSerGluThrProLeuSerPro 340
DB 1547 ATGAACCTGCTCCCTGCTGCTCAGAGGGTGAAGAAAGCGAGAGCCCTTCACTGCTC 1606
QY 340 roCysAlaAspSerAlaThrSerGluAlaIleGlyIleArgProValHisProAlaSer 360
DB 1607 CCTGTGCCACTCTGCTACTCAGAGGCGAGGAGCAAAAGCCCGTGAATCCCGCAAG 1666
QY 360 aIProSerValProSerValGlnProSerLeuLeuGluIleAspSerProLeuValGlnAla 380
DB 1667 TGCCAGCTGCTCCAGCGTGCAGCCGCTGTTAGAGACAGCCGCTGGTACAGGCGC 1726
QY 380 euSerGlyLeuHisIleSerAlaAspArgLeuGluIleAspSerAsnIleLeuValArgSer 400
DB 1727 TCAGTGGCTGCACCTGCTCGCAGAGAGCTGGAAGACGAAATAGCTCAAGCGTTCT 1786
QY 400 HeSerLeuAspIleLeuSerValSerTyriSerAlaSerMetAlaIleSerLeuHisGly 420
DB 1787 TCTCTGTGATTCAAATAGTTTCAATATTCAGCCAGCACTGACAGATCTTACATGCT 1846
QY 420 HeSerSerSerGlnAspAlaLeuGluTyriTyriLeuProSerThrThrLeuAspGlyThr 440
DB 1847 TCTCTCATCAAGAGTGTGTTGAAATCTCAAAACCTTCCACTCTGATGGAGCA 1906
QY 440 hNileLeuCysGlnPheSerProValGlnIleLeuSerGluGlnThrProGluThrSer 460
DB 1907 ACAAGCTATGCAAGTTCTCCCTGTTCAAGAACTATCGAGAGACTCCGAAACCAAG 1966
QY 460 roAspIleGluGluAlaSerIleProIleGlyLeuGlnIleThrAlaArgProSerAspSer 480
DB 1967 CTGATTAAGAGAGAACCGACATCCCAAGAACTGAGACCCGACAGCTTTCAGAGAGC 2026
QY 480 InSerIleArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
DB 2027 AAGCAGAGGATTCATTCGTCAGAACCGAGACGAGTGGACCCGCCAGAGTCCCTTT 2086
QY 500 euSerProLeuHisArgSerGlySerValGluAspAsnTyriArgThrPheLeuPheG 520
DB 2087 TATCTCCATGTCATGAGTGGAGGTGAGAGCAATTAACCAACACCACTTCTTTTTCG 2146
QY 520 ILeuSerThrSerGlnGlnHisIleuThrIleValSerAlaGlyLeuGlyIleValGlyTyr 540
DB 2147 GCTTTTCCACCGCAGCAGCACTTACGAACTGCTGGCCCTTAAAGGCTGAC 2206
QY 540 IeSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyriTyriPhe 560

Db 2207 ACTCGGATATCTTGCGCCCGACGACTTACCCCTCTCGACGAGCGATGATTTTG 2266
Qy 560 lathrguserSerHisPheTySerAlaSerAlaIleTyrglyGlySerAlaSerTyrs 560
Db 2267 CCACAGATCTCTACACTTCTACTCTCTCAGCATCTACGAGGAGGACAGTTACT 2326
Qy 580 ealalyrSerCySerGlnLeuProThrCysGlyAspGlnValTySerValArgArg 600
Db 2327 CTGCTACAGCTGAGCAGCTGCGCCACTTGCGGAGACCAAGCTATTTCTGTGCGCAGGC 2386
Qy 600 rgtGlnySProSerAspArgAlaAspSerArgArgSerTyphsglyGlySerProphg 620
Db 2387 GGCAGAGCCAGAGTGAAGAGCTGACTCGCGGAGCTGCATGAAGAGCCCTTTG 2446
Qy 620 lutySGlnPheTyArgArgSerCysGlnMetGlnPheGlyGlyuserIleMetSerGln 640
Db 2447 AAAAGCAGTTTAAACGAGAACTGCCAAATGGAAATTTGAGAGAGCATCATCTAGAGA 2506
Qy 640 snArgSerArgGlnGlnLeuGlyLyValGlySerGlnSerSerPheSerGlySerMetC 660
Db 2507 ACAGGTCAACGGAAAGAGCTGGGAAAGTGGGAGTCACTTCTTTCGGGAGAGCATG 2566
Qy 660 lullelleglyValSer 665
Db 2567 AAATCATGTGAGGTCTCC 2583

RESULT 27

ABV26934 standard; cDNA; 5145 BP.
XX

ABV26934;
XX

16-SEP-2002 (first entry)
XX

Human prostate expression marker cDNA 26925.
XX

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX

pharmacogenomic marker; gene; ss.
XX

Homo sapiens.
XX

MO200160860-A2.
XX

23-AUG-2001.
XX

20-FEB-2001; 2001WO-US005171.
XX

17-FEB-2000; 2000US-0183119P.
XX

16-MAR-2000; 2000US-0189862P.
XX

25-MAY-2000; 2000US-0207454P.
XX

09-JUN-2000; 2000US-021314P.
XX

18-JUL-2000; 2000US-0219007P.
XX

13-DEC-2000; 2000US-0255281P.
XX

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX

Schlegel R, Endege WO, Monahan JE;
XX

WPI; 2001-662795/76.
XX

Novel isolated nucleic acid molecule associated with cancerous state of
XX

prostate cells and correlating with presence of prostate cancer, useful
XX

for detecting presence of prostate cancer, stage of prostate cancer.
XX

Claim 1; Page 5453-5454; 11750pp; English.
XX

CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	0	Length:	5145
Score:	472.00	Matches:	663
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	70.98%	Indels:	4
DB:	5	Gaps:	0

US-10-029-345A-109 (1-665) x ABV26934 (1-5145)

Qy 1 MetAlaHisGlyMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGln 20
Db 589 ATGGCCATGAGATGATGAACTCAAAATGTCTGAGAGGTGGCTCTCTGGAA 648
Qy 21 SerGlyThrGlnValLeuLeuIleAspSerArgProPheValGluTyArgThrSer 40
Db 649 AGTGAAGCGAAAAAGTCTCTAATGATGACCGGCATTTGTGAATACATCATCC 708
Qy 41 HisIleLeuGlnAlaIleAsnIleAsnCysSerIysLeuMetIyArgArgLeuGln 60
Db 709 CACATTTTGAAAGCATTAAATATCACTCCAAAGCTTATGAAGCCGAAGTTCACAG 768
Qy 61 AspIysValLeuIleThrGlnLeuIleGlnHisSerAlaIySHslsValAspIleAsp 80
Db 769 GACAAAGTGTAAATTAACAGACTCATCCAGCATTCAGGAAACATTAAGTTGACATTGAT 828
Qy 81 CysSerGlnIyValValIyTyArgPqlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 829 TGCACTCAGAGGTTGATGATTAAGATCAAAAGTCCCAAGATGTGCTCTCTTCA 888
Qy 101 AspCysPheLeuThrValLeuLeuGlyLyValLeuGlnIySerPheAsnSerValHisLeu 120
Db 889 GACTGTTTCTCAGCTGATCTTCTGCTGTAACCTGGAAGAGAGCTTCACTGTTCACCTG 948
Qy 121 LeuAlaGlyIyPheAlaGlyPheSerArgCysPheProGlyLeuCysGlnIyLySer 140
Db 949 CTTCAGAGTGGGTTTGTCTGAGTTCTCTGTTGTTTCCCTGCTCTGTGAAGAAATTC 1008
Qy 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 1009 ACTCTAGTCCCTACCTGATTTCTCAGCTTGTCTTACCTGTGCAACATTTGGGCCAAC 1068
Qy 161 ArgIleLeuProAsnLeuTyrlleuGlyCysGlnArgAspValLeuLeuIyGlyLeu 180
Db 1069 CGAATCTTCCCAATCTTAACTTGTGCTGCCAGAGATCTCTCAACAGAGAGCTGAT- 1127
Qy 181 -GlnGlnAsnGlyIleGlyTyrlValLeuAsnAlaSerTyrl-ThrCysProIyProAsp 200
Db 1128 GCAGCAAGATGGAGTTGTTATGTTAATGCAAGAA-TACTGTGCCAAACCTGACT 1186
Qy 200 helLeProGlyuserHisPheLeuArgValProValAsnAspSerPheCysGlyIyIle 220
Db 1187 TTTATCCCGAGTCTCAATTTCTGCTGCTGTGATGACACATTTTGGAAATAATTT 1246
Qy 220 eupProTrPheLeuAspIySerValAspPheIleGlnIyAspIyAspIyCysGly 240
Db 1247 TGCCTGCTGTGCAAAATCAAGATTCATTTAGAAAGCAAAAGCTCCAAAGATGCTG 1306
Qy 240 alleuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlle 260
Db 1307 TTTTAAAGCACTGTTTAAAGTGGATCTCCGCTCCGCCACACATCGCTTACATCA 1366
Qy 260 etIyArgMetAspMetSerLeuAspGlnAlaTyArgPheValIyGlyLyArgProT 280
Db 1367 TGAAGAGATGACATGTCTTTAGATGAAGCTTACGATTTGCAAAAGAAAGACCTTA 1426

QY 280 hrllserProaenPheAenPheLeuGlyGlnLeuLeuSerPyrGlyLysLysLysLysLysLys 300
 Db 1427 CTATATCTCCAACTCAATTTCTGAGCACTCTGAGCATATGGAAGAAATTAAGA 1486
 QY 300 englnThrGlyAlaSerGlyProLysSerLysLysLysLysLysLysLysLysLysLys 320
 Db 1487 ACCAGACTGGAGCATCAGGGCCAAAGAGCAATCAAGCTGCTGAGCACTGGAGAACCA 1546
 QY 320 englnProValProAlaValSerGlyGlyGlnLysSerGlyLysLysLysLysLysLys 340
 Db 1547 ATGAACTGCTCCCTGCTGCTCAGAGGCTGAGCAAGAAAGAGAGAGAGAGAGAGAGAG 1606
 QY 340 rocCyAlaAapSerAlaThrSerGlyAlaAlaGlyGlnPyrProValHisProAlaSerV 360
 Db 1607 CCTGTGCCAGCTCTGCTACTCTCAGAGCAGAGCAAGAGAGAGAGAGAGAGAGAGAGAG 1666
 QY 360 aLProSerValProSerValGlnProSerLeuLeuGlnAapSerProLeuValGlnAla 380
 Db 1667 TGCCCAAGCGTCCCAAGCGTCCGCTGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1726
 QY 380 euSerGlyLeuHisLeuSerAlaAapArgLeuGlnAapSerAenLysLysLysLysLys 400
 Db 1727 TCAGTGGGCTGCACCTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1786
 QY 400 hseSerLeuAapLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyP 420
 Db 1787 TCTCTGTGAATACAAATCATGTTCAATTCACGAGCATGAGAGATCTTTCATGAGTCT 1846
 QY 420 hseSerSerGlyAapAlaLeuGlyLysTyrLysProSerThrThrLeuAapGlyThra 440
 Db 1847 TCTCCCATCAGAAAGATGCTTGGAAATCATCAAACTTCCACTACTCTGGATGGAGCA 1906
 QY 440 enLysLeuCyGlnPheSerProValGlnGlnLeuSerGlyGlnThrProGlyLysThr 460
 Db 1907 ACAAGCTATGCCAGTTCTCCCTGTTACAGGAATATCGAGAGAGAGAGAGAGAGAGAG 1966
 QY 460 roaApyrGlyGlnGlyAlaSerLysProLysLysLysLysLysLysLysLysLysLys 480
 Db 1967 CTGATTAAG 2026
 QY 480 lnsSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
 Db 2027 AAGAGCAAGGATTCGATTCGGTTCAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2086
 QY 500 euSerProLeuHisArgSerGlySerValGlnAapAenLysLysLysLysLysLysLys 520
 Db 2087 TATCTCAGCTGATCGAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2146
 QY 520 lylLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLysLysLysLysLysLys 540
 Db 2147 GCGTTTCCACAGCCAG 2206
 QY 540 lsserApyrLysLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheA 560
 Db 2207 ACTCGAATATCTTGAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2266
 QY 560 lathrGlnSerSerHisPheTyrSerAlaSerAlaLysLysLysLysLysLysLysLys 580
 Db 2267 CCACAGAGCTCTCAGACTTCTACTGCTCAGCCATTCAGAGAGAGAGAGAGAGAGAGAGAG 2326
 QY 580 eraLysLysSerCySerGlnLeuProThrCySerGlyAapGlnValLysSerValArgArg 600
 Db 2327 CTGCCCTACAGCTGACAG 2386
 QY 600 rglLysPyrProSerAapArgAlaAapSerArgArgSerTyrPheGlnGlnLysSerProPhe 620
 Db 2387 GCGAG 2446
 QY 620 lylLysGlnPheLysArgArgSerCyGlnMetGlnPheGlyGlnSerLysLysSerGln 640
 Db 2447 AAAAGAGATTAAACAG 2506

QY 640 sArGSerArgGlnGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetC 660
 Db 2507 AAGGTCACGGAG 2566
 QY 660 lnllelGlnValSer 665
 Db 2567 AAATCATTGAGGTCCTCC 2583
 RESULT 28
 AAF30479
 ID AAF30479 standard; cDNA; 2118 BP.
 XX
 AC AAF30479;
 XX
 DT 29-MAY-2001 (first entry)
 XX
 DE Human protein phosphatase and kinase protein-4 cDNA 1234795CBL.
 XX
 KW Protein phosphatase and kinase protein; PPHKP-4; human;
 KW gastrointestinal disorder; immune system disorder; neurological disorder;
 KW cell proliferative disorder; cancer; diagnosis; therapy; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 25..2025
 FT /tag= a
 FT misc_feature 163..222
 FT /tag= b
 FT /note= "unique fragment"
 FT misc_feature 1081..1140
 FT /tag= c
 FT /note= "unique fragment"
 XX
 PN WO200120004-A2.
 XX
 PD 22-MAR-2001.
 XX
 PF 14-SEP-2000; 2000MO-US025515.
 XX
 PR 15-SEP-1999; 99US-0154141P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y;
 PI Lu DAM;
 DR WPI: 2001-244811/25.
 DR P-PSDB; AAB20325.
 PS Claim 5; Page 97-98; 103pp; English.
 XX
 CC The present sequence is that of cDNA encoding novel human protein
 CC phosphatase and kinase protein PPHKP-4 (see AAB20325). The cDNA was
 CC initially identified in Incyte Clone ID No. 1234795CBL, from a foetal
 CC lung tissue library. Tissues that express PPHKP-4 (as a fraction of total
 CC tissues expressing PPHKP-4) include gastrointestinal (0.385),
 CC cardiovascular (0.154), nervous (0.154) and reproductive (0.154).
 CC Diseases or conditions associated with tissues expressing PPHKP-4 (as a
 CC fraction of total tissues expressing PPHKP-4) include cancer (0.692),
 CC inflammation or trauma (0.308) and cell proliferation (0.231). The
 CC encoded protein shows homology to mouse neuronal tyrosine threonine
 CC phosphatase 1. The invention provides human PPHKP-1 to -11 polypeptides
 CC (see AAB20322-32) and polynucleotides (see AAF30475-86). It also provides
 CC expression vectors, host cells, antibodies, agonists and antagonists, as
 CC well as methods for diagnosing, treating or preventing disorders
 CC associated with expression of PPHKP, including gastrointestinal
 CC disorders, immune system disorders, neurological disorders and cell
 CC proliferative disorders, including cancer

XX Sequence 2118 BP; 550 A; 552 C; 506 G; 510 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 0 Length: 2118
Score: 439.00 Matches: 663
Percent Similarity: 98.96% Conservative: 0
Best Local Similarity: 98.96% Mismatches: 2
Query Match: 66.02% Indels: 7
DB: 4 Gaps: 0
US-10-029-345a-109 (1-665) x AAF30479 (1-2118)
QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 25 ATGGCCCATGAGTATGATGGAACTCAAAATTTGTAAGAGAGTGGTGGCTCTGCTGAA 84
QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db 85 AGTGAACCGGAAAGGTGCTGTAATTGATAGCCGGCCATTGTGGAAATACATACATCC 144
QY 41 HisIleLeuGluAlaIleAsnIleAsnGlySerIleLeuMetLysArgArgLeuGlnGln 60
Db 145 CACATTTTGGAAAGCCATTAATATCACTGCTCCAACTTAATGAAGGAAAGTTGCAACAG 204
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db 205 GACAAAGTGTATTTACAGAGCTCATCCAGCATTCAGCAACATTAAGTTGACATTTGAT 264
QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 265 TGGAGCAGAAAGGTGTAGTTTACGATCAAGAGCTCCCAAGATGTTGCTCTCTCTTCA 324
QY 101 AspCysPheLeuThrValLeuLeuGluLysLeuGluLysSerPheAsnSerValHisLeu 120
Db 325 GACTGTTTCTCACTGACTCTCGGATTAACGTGGAAGAAGCTTCAACTCTGTTCACTG 384
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGlnGlyLysSer 140
Db 385 CTGTGAGGTGGGTTTGTGATCTCTGTTTCCCTGGCTCTGTCAAGAAATTC 444
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 445 ACTCTAGTCCCTACCGCATTTCTCAGCTTCTTAACCTTGAACCAACTGGGCCAACCC 504
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnLysGluLeu 180
Db 505 CCAATTTCTTCCCAATCTTATCTTGCTGCCAGCAGCATGCTCTCAACAAAGAGCTGAT- 563
QY 181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProLysProAsp 200
Db 564 GCAGCAGATGGGATTTGTTATGTGTTAAATGCCAGCA-TACGTGTCCAAAGCTGACT 622
QY 200 IleIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle 220
Db 623 TTATCCCCCAGTCTCATTTCTCGCGTGCCCTGTAAATGACAGCTTTGTGAGAAATTT 682
QY 220 euProTribLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCysV 240
Db 683 TGCCGTGTGGTGGCAAAATCAGTAATTTTCATTGAGAAAGCAAAAGCCTCCATATGATG 742
QY 240 alLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleM 260
Db 743 TTCTAGTGCACATGTTAGCTGGGATCTCCGCTCCGCCACCATGCTATGCTATACATCA 802
QY 260 eLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProT 280
Db 803 TGAAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAGAAAGACCTTA 862
QY 280 hrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLysIleLysA 300
Db 863 CTATATCTCCAAACTTCAATTTTCTGGGCCAATCTCTGAGACTATGAGAAAGATTTAAGA 922

QY 300 enGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGluLysProA 320
Db 923 ACCAGACTGAGACATCAGGCGCAAGAGCAAACTAAGCTGTGTGACCTGGAGAAAGCCAA 982
QY 320 enGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProp 340
Db 983 ATGAACCTGTCTCTGCTGTCTGTCAGAGGGTGGACAAAGAAAGAGAGCGCCCTCATTCAC 1042
QY 340 rOCyAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerV 360
Db 1043 CCTGTGCGGACTTCTTACCTTACGAGGAGCAGGCAAGAGCCCTGTGATCCCGCAGCG 1102
QY 360 alProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAla 380
Db 1103 TGCCAGAGTCCCGCTGAGCCGTGCTGTTAAGAGACAGCCGCTGGATCAGGGCG 1162
QY 380 euSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerP 400
Db 1163 TCAGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCTTCTCT 1222
QY 400 heSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyP 420
Db 1223 TCTCTGTGATATCAATCAGTTTCATATTCAGCCAGCATGCGACATCTTACATGCT 1282
QY 420 heSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThra 440
Db 1283 TCTCTCTATCAGAAAGTCTTTGGAATCTTCAAACTTCCACTACTTGTGATGGAGACA 1342
QY 440 snLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSer- 459
Db 1343 ACAAGCTATGCGCAGTTCTCCCTGTTGAGAACTATCGAGACAGCTCCGAAACAGATT 1402
QY 460 -ProAspLysGlnGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSe 479
Db 1403 CCTGATTAAGAGAGAGAGCAGACATCCCAAGAACTGAGACCCGAGGCTTACAGACG 1462
QY 479 rGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLe 499
Db 1463 CCAAGCAGAGGATTTGATTTGCTGTCAGAACCAAGCAGACGACCGCCAGAGTCTCT 1522
QY 499 uLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPh 519
Db 1523 TTTATCTTCACATGCATCGAAGTGGAGGTGAGAGACATTAACACACAGCTTCTTTT 1582
QY 519 eGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGly-LeuLysGlyT 539
Db 1583 CGGCTTTTCCACAGCAGCAGCAGCACTTCAAGAGTGTGCTGGCTGGCCCTTTAAGGCT 1642
QY 539 rPhIleSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTribLysP 559
Db 1643 GGCACCTGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGATATT 1702
QY 559 heaIaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerT 579
Db 1703 TTGGCAGAGAGTCTTCACTTACTGCTGCTCAGCCATCTAGGAGGAGCAGTCCAGTT 1762
QY 579 rYrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArg 599
Db 1763 ACTGTGCTTACAGCTGACGCGACGCTGCCACTTGGGAGAGCAAGTATTTCTGTGGCA 1822
QY 599 rGATrGlnLysProSerAspArgAlaAspSerArgAspSerTribIleGluLysSerProp 619
Db 1823 GCGCGGAGAAAGCCAAAGTACAGAGCTACTCGCGCGGAGTGGCATGAAGAGAGCCCT 1882
QY 619 heGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerG 639
Db 1883 TTGAAAGCAGTTTAAACGAGAAAGCTGCCAAATGCAATTTGAGAGAGCATATGTCAG 1942
QY 639 lAsnArgSerArgGluGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerM 659
Db 1943 AGAACAGGTACCGGAAAGAGCTGGGAGAAAGTGGGAGTCACTAGCTTTTTCGGGAGCA 2002
QY 659 etGluIleIleGluValSer 665

Db 2003 TGGAAATCATTTGAGTCTCC 2022

RESULT 29

ACC60521
ID ACC60521 standard; cDNA; 5111 BP.

AC ACC60521;

DT 19-JUN-2003 (first entry)

XX Polynucleotide relating to the invention SEQ ID NO: 41.

XX Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic;
XX antipsoriatic; cardiant; cyrostatic; gene therapy; liver disease;
XX proliferative disorder; renal failure; cardiovascular disorder;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.

OS Homo sapiens.

XX WO200257460-A2.

XX 25-JUL-2002.

XX 20-DEC-2001; 2001WO-US050459.

XX 20-DEC-2000; 2000US-0256868P.

XX 30-MAR-2001; 2001US-0280186P.

XX 01-MAY-2001; 2001US-0287735P.

XX 05-JUN-2001; 2001US-0295848P.

XX 25-JUN-2001; 2001US-0300465P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
XX Siemere N, Bol D, Schieven G, Finger J, Todderud CG, Basolino D;
XX Krystek S, Mcatee P, Suchard S, Banas D;

XX WPI, 2002-599721/64.

XX P-PSDB; ABR52352.

XX Novel polynucleotides encoding human phosphatase polypeptides useful in
XX the prevention or treatment of e.g. proliferative and cardiovascular
XX disorders.

XX Claim 1; Fig 5; 801pp; English.

XX The invention relates to a novel isolated nucleic acid comprising a
XX polynucleotide having a nucleotide sequence selected from 40
XX polynucleotides fully defined in the specification. The polynucleotide of
XX the invention has antiproliferative, hepatotropic, nephrotropic,
XX antiarthritic, antipsoriatic, cardiant, and cyrostatic activity. The
XX polynucleotide may have a use in gene therapy. A polynucleotide or
XX polypeptide of the invention is useful for preventing, treating or
XX ameliorating a medical condition, e.g. a proliferative disorder. They are
XX also useful for treating e.g. liver disease, renal failure, immunological
XX disorders including arthritis and psoriasis, cardiovascular disorders
XX such as congenital heart defects and congestive heart failure, and
XX cancer. A method of the invention is useful for diagnosing a pathological
XX condition or susceptibility to a condition in a subject. The present
XX sequence is used in the exemplification of the invention

XX Sequence 5111 BP, 1413 A, 1112 C, 1232 G, 1354 T, 0 U, 0 Other;

Alignment Scores:

Pred. No.: 0
Score: 350.00
Percent Similarity: 98.808
Best Local Similarity: 98.808
Query Match: 52.634
DB: 6

Length: 5111

Matches: 661

Conservative: 0

Mismatches: 4

Indels: 8

Gaps: 0

US-10-029-345A-109 (1-665) x ACC60521 (1-5111)

QY	1	MeAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuGlu	20
DB	470	ATGGCCCATGAGATGATTGAACTCAATTTGACTGAGAGGTGGTGGCTGCGGAA	529
QY	21	SerGlyThrGluValLeuLeuIleAspSerArgProPheValGlnTyrAenThrSer	40
DB	530	AGTGAACCGGAAAAAGTGGCTGCTAATTGATAGCCGCCATTGGAAATCAATATATCC	569
QY	41	HisIleLeuGluAlaIleAsnIleAsnCySerIleValLeuMetIleValArgLeuGln	60
DB	590	CACATTTTGAAGCCATATATATCACTGCTCCAACTTATGAAAGGAAGTTGCAACG	649
QY	61	AspIleValLeuIleThrGluLeuIleGlnHisSerAlaIleValIleAspIleAsp	80
DB	650	GACAAAGTGTATATACAGAGCTCAACGCAATTCGACCAAACTATAGGTGACATTTAT	709
QY	81	CySerGlnIleValIleValIleAspGlnSerSerGlnAspValIleSerIleSerSer	100
DB	710	TGCAGTCAGAGGTTGTATTGATTCATCAATCAAGCTCCCAAGATGTTGCTCTCTTCA	769
QY	101	AspCySerPheLeuThrValLeuLeuGlyIleValLeuGlySerPheAsnSerValIle	120
DB	770	GACTGTTTCTCAGCTACTTCTGGGTAACTGAGAGAGGTTCAACTCTGTCACCTG	829
QY	121	LeuAlaGlyIlePheAlaGluPheSerArgCySerPheProGlyLeuCySerGlyIle	140
DB	830	CTTGACAGTGGGTTGCTGAGTCTCTCTGTTTCCCTGCGGCTCTGTAAGGAAAAATCC	889
QY	141	ThrIleValProThrCySerIleSerGlnProCySerProValAlaAsnIleGlyPro	160
DB	890	ACTTATGCTTCACTGCTCAATTTCTCAGCTTCTGCTTCACTGCTTCACTGCTTCA	949
QY	161	ArgIleLeuProAsnLeuTyrLeuGlyCySerArgAspValLeuAsnIleGlyLeu	180
DB	950	CGAATTTCTCCAAATCTTATATCTTGGCTGCGGAGAGATGCTCTCAACAGAGCTGAT	1008
QY	181	GlnGlnAsnGlyIleGlyTyrValIleAsnAlaSerTyr-ThrCySerProIleAsp	200
DB	1009	GCAGCAGATGGAGATTGTTATGTGTTAATCCAGCAA-TACCTGTCACAAAGCTCGACT	1067
QY	200	HeIleProGluSerHisPheLeuArgValProValAsnAspSerPheCySerIleLeu	220
DB	1068	TATTCCTCCGAGTCTCAATTTCTGCGGTGCTGTTGATACAGCTTTTGGAAAAATTT	1127
QY	220	euProThrLeuAspIleSerValAspPheIleGlyIleValIleValIleSerAsnGly	240
DB	1128	TGCCGTGGTTGACAAATCAGTAGATTTCAATTGAGAAACAAAGCTCCATGATGTTG	1187
QY	240	AlleuValHisCySerLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle	260
DB	1188	TTCTGTGTCAGCTTTAGCTGGGAGTCTCCGCTCCGCCACATCGCTATCGCTATCATCA	1247
QY	260	eIleValArgMetAspMetSerIleValAsnGluAlaTyrArgPheValIleValIle	280
DB	1248	TGAAGAGATGAGACATGTTTATGATGAGAGCTTACGATTGGAAGAAAAAGACCTTA	1307
QY	280	HisIleSerProAsnPheAsnPheLeuGlyIleLeuLeuAsp-TyrGlnIleValIle	299
DB	1308	CTATATCTTCAAACTTCAATTTTTCGGGCAACTCTGCGC-CTATGAGAGAAATTAAG	1366
QY	300	AenGlnThrGlyAlaSerGlyProIleSerIleValLeuLeuHis-LeuGlyIleAsp	319
DB	1367	AACCAAGACTGAGCATCAGGAGCAAAAGCAAACTCAACTGCTGCGC-CCTGAGAGAGCC	1425
QY	319	OhAnGluProValProAlaValIleSerGluIleGlyIleGlnIleSerGluThrProLeu	339
DB	1426	AAATGAACCTGCTCCGCTGCTTCAAGGGGTGACGAAAGAGAGCCCTCTGCTGTC	1485
QY	339	oProCyAlaAspSerAlaThrSerGlnAlaIleGlyIleGlnIleGlnProValHisPro	359
DB	1486	ACCCTGTGCGCACTGCTGCTACTCAGAGGACAGACAGAAAGCCCGTGTCACTCCGCG	1545


```

QY      359      fValProSerValProSerValGlnProSerLeuGluAspSerProLeuValGlnAl 379
DB      1546      GGTCCCAAGCGTCCAGCGCGCGTGTAGAGACAGCCCGCTGTACAGGC 1605
QY      379      aLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuValAspSe 399
DB      1606      GCTCAGTGGGCTGCACTGTCTCCGACAGACGCTCGAAGACAGCAATAGCTCAAGCCTTC 1665
QY      399      rPheSerLeuAspLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGly 419
DB      1666      CTCTCTCTCGAATATCAAAATCAGTTTATTCATTCAGCAGCATGAGCATCTTACATGG 1725
QY      419      yPheSerSerSerGluAspAlaLeuGluLysTyrTyrLysProSerThrThrLeuAspGly 439
DB      1726      CTCTCTCTCATCAGAAATGCTTTGAAATCTCAAAACCTTCCACTACTCTGATGGAGAC 1785
QY      439      rAsnLysLeuGluGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSe 459
DB      1786      CAACAAGCTATGCGCAGTTCTCCCTGTTCAGGAACATATCGAGCAGACTCCGAAACAG 1845
QY      459      rProAspLysGluGluAlaSerLysLeuLysLysLeuGlnThrAlaArgProSerAspSe 479
DB      1846      TCCTGATAGAGAGAGAGCCAGCATCCCAAGAAAGCTGACAGCCGACGAGCTTCAACAG 1905
QY      479      rGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLe 499
DB      1906      CCAGAGAGACCGATTGCTCGGTCAAGAACACAGACAGAGTGGACGCGCCAGAGGTCCCT 1965
QY      499      uLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeu 519
DB      1966      TTATATCTCCACTGATCGAAGTGGAGCGTGGAGGACAATATCACACAGGCTTCTTTT 2025
QY      519      eGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyr 539
DB      2026      CGGCTTTTCCACAGCAGCAGACGACCTCAGAAAGTCTGCTGGCGCTTAAAGGCGTGG 2085
QY      539      pHisSerAspLysLeuAlaProGlnThrSerThrProSerLeuThrSerSerTriPyr 559
DB      2086      GCACCTCGAATATCTTGCCCCCCAGACCTCTACCCCTTCCCTGACAGACGCTGATATT 2145
QY      559      eAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyr 579
DB      2146      TGGCAGAGAGTCTCTCACTTCTACTCTCTCAGCAGCATCTACGAGGAGGAGCCAGTTA 2205
QY      579      rSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgAr 599
DB      2206      CTCTGCTTCAAGCTGACGACGCTGCTCCACTTGGGAGACCAAGTCTATTCTGTGCGCAG 2265
QY      599      gArgGlnLysProSerAspArgAlaAspSerArgArgSerTPHISGluGluSerProPh 619
DB      2266      GCGGCGAGAGCCAGACGACAGAGCTGACTCGCGCGGAGCTGCATGAAGAGACCCCTT 2325
QY      619      eGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerTleMetSerGly 639
DB      2326      TGAATAAGCAGTTTAAACGCAAGAGCTGCCAAATGGAATTTGAGAGAGCATCATGTCA 2385
QY      639      uAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMe 659
DB      2386      GAAACAGGTACCGGAAGAGCTGGGAAAGTGGGAGAGTCACTTCTTGGGAGAGCAT 2445
QY      659      tGluIleIleGluValSer 665
DB      2446      GGAATATCATTTGAGGTCTCC 2464

```

RESULT 30
ACC60560
ID ACC60560 standard; cDNA; 2756 BP.

XX ACC60560;
XX
DT 19-JUN-2003 (first entry)
XX

```

DE Polynucleotide relating to the invention SEQ ID NO: 113.
XX
XX Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic;
XX antipsoriatic; cardiant; cytosaric; gene therapy; liver disease;
XX proliferative disorder; renal failure; cardiovascular disorder;
XX immunological disorder; arthritis; psoriasis; congenital heart defect;
XX congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
OS Mus musculus.
XX
XX WO200257460-A2.
XX
XX 25-JUL-2002.
XX
XX 20-DEC-2001; 2001WO-US050459.
XX
XX 20-DEC-2000; 2000US-0256868P.
XX 30-MAR-2001; 2001US-0280186P.
XX 01-MAY-2001; 2001US-0287735P.
XX 05-JUN-2001; 2001US-0295848P.
XX 25-JUN-2001; 2001US-0300465P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L,
XX Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D,
XX Krystek S, Mcatee P, Suchard S, Banae D;
XX
XX WPI; 2002-599721/64.
XX
XX P-PSDB; ABR52385.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
XX the prevention or treatment of e.g. proliferative and cardiovascular
XX disorders.
XX
XX Disclosure; Fig 16; 801pp; English.
XX
XX
XX The invention relates to a novel isolated nucleic acid comprising a
XX polynucleotide having a nucleotide sequence selected from 40
XX polynucleotides fully defined in the specification. The polynucleotide of
XX the invention has antiproliferative, hepatotropic, nephrotropic,
XX antiarthritic, antipsoriatic, cardiant, and cytosaric activity. The
XX polynucleotide may have a use in gene therapy. A polynucleotide or
XX polypeptide of the invention is useful for preventing, treating or
XX ameliorating a medical condition, e.g. a proliferative disorder. They are
XX also useful for treating e.g. liver disease, renal failure, immunological
XX disorders including arthritis and psoriasis, cardiovascular disorders
XX such as congenital heart defects and congestive heart failure, and
XX cancer. A method of the invention is useful for diagnosing a pathological
XX condition or susceptibility to a condition in a subject. The present
XX sequence is used in the exemplification of the invention
XX
XX
XX Sequence 2756 BP; 714 A; 675 C; 664 G; 703 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1 49e-52 Length: 2756
XX Score: 64.00 Matches: 93
XX Percent Similarity: 97.89% Conservative: 0
XX Best Local Similarity: 97.89% Mismatches: 1
XX Query Match: 9.62% Indels: 2
XX DB: 6 Gaps: 0
XX
XX US-10-029-345A-109 (1-665) x ACC60560 (1-2756)
XX
XX QY 194 ThrCysProLysProAspPheIleProGluSerHisPheLeuArgValProValAsnAsp 213
XX Db 948 ACCTGTCGAAGCCTGACTTACATACCGAATCTCACTTCTGCGAGTGGCTGTGAATGAC 1007
XX
XX QY 214 SerPheCysGluLysIleLeuProTTPLeuAspLysSerValAspPheIleGluLysAla 233
XX Db 1008 AGCTTTTGTGAAGAAATCTTACCATGTTGGACAAAGTCTGTGATTTTCAATTGGAAGAACA 1067
XX
XX QY 234 LysAlaSerAsnGlyCysVal-LeuValHisCysLeuAlaGlyLysSerArgSerAlaThr 253

```



```

DB: 1068 AAGCTTCAGTGGCTGTGCTTATC-CAGTCTAGTGGGATCTCGTCCGCCAC 1126
QY: 253 rllleallleallatYrllleWetlysaRgWetlaPctSerleuAepGlualatYraRph 273
DB: 1127 TATTCCTATGCTCATCATCATGAGATGAGACATGCTCTAGATGAGGCTTACAGATT 1186
QY: 273 eVallysGlylysaRgProThrllleSerProAenPheanPhe 287
DB: 1187 TGTGAAGAAAGAAAGACCTACTATATCTCCGAATTTAATTTT 1229

RESULT 31
ADE07547
ID ADE07547 standard; DNA; 1917 BP.
AC ADE07547;
XX
XX
XX 29-JAN-2004 (first entry)
XX
XX
XX Novel coding sequence (useful for identifying genetic disorders) #613.
XX
XX novel gene; novel protein; tissue marker; molecular weight marker;
XX chromosome marker; genetic disorder; gene; ds.
XX
XX Unidentified.
XX
XX WO2003054152-A2.
XX
XX 03-JUL-2003.
XX
XX
XX 10-DEC-2002; 2002MO-US039555.
XX
XX
XX 10-DEC-2001; 2001US-0339739P.
XX
XX 11-DEC-2001; 2001US-0339453P.
XX
XX 14-MAR-2002; 2002US-0365091P.
XX
XX 14-MAR-2002; 2002US-0365384P.
XX
XX 12-APR-2002; 2002US-0372381P.
XX
XX 12-APR-2002; 2002US-0372615P.
XX
XX 22-APR-2002; 2002US-00128558.
XX
XX 24-APR-2002; 2002US-0376045P.
XX
XX
XX (HYSE-) HYSEQ INC.
XX
XX
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
XX Ghosh W, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
XX Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX
XX WPI; 2003-569235/53.
XX
XX P-PSDB; ADE08458.
XX
XX
XX New polynucleotides, useful for expressing recombinant proteins for
XX analysis, characterization or therapeutic use, or as markers for tissues
XX in which the corresponding protein is preferentially expressed.
XX
XX
XX Claim 1; SEQ ID NO 613; 1177bp; English.
XX
XX
XX The invention comprises the amino acid and coding sequences of novel
XX proteins. The DNA and protein sequences of the invention are useful as:
XX markers for tissues in which the corresponding protein is preferentially
XX expressed; as molecular weight markers on gels; as chromosome markers or
XX tags; to identify chromosomes or to map related gene positions; and to
XX compare with endogenous DNA sequences in patients to identify potential
XX genetic disorders. The present DNA sequence represents a gene of the
XX invention.
XX
XX
XX Sequence 1917 BP; 305 A; 724 C; 613 G; 275 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.86e-09 Length: 1917
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.01% Indels: 0

```

```

DB: 9 Gaps: 0
US-10-029-345A-109 (1-665) x ADE07547 (1-1917)
QY: 242 VALHISQYsleuAlGlyllleSerArgSerAlaThrllleAlaIleAlaTyrllleWetly 261
DB: 733 GTCCACTGCTGCTGGCGCATCTCCGCTCTGCACCATGCGCATGCGCTACATCATGAG 792

RESULT 32
ACA64956
ID ACA64956 standard; DNA; 2377 BP.
AC ACA64956;
XX
XX
XX 27-JUN-2003 (first entry)
XX
XX
XX Human protein tyrosine phosphatase DNA corresponding to U27193.
XX
XX
XX Human; chronic inflammatory joint disease; infection; tumour;
XX antiinflammatory; cytosratic; antirheptic; antirheumatic;
XX immunosuppressive; gene therapy; etiological pathogenicity; ds.
XX
XX Homo sapiens.
XX
XX DE10127572-A1.
XX
XX
XX 05-DEC-2002.
XX
XX
XX 30-MAY-2001; 2001DE-01027572.
XX
XX
XX 30-MAY-2001; 2001DE-01027572.
XX
XX
XX (PATH-) PATHOARRAY GMBH.
XX
XX
XX Haepul T, Ungethuem U, Blaess S;
XX
XX WPI; 2003-240797/24.
XX
XX
XX Reagents for diagnosis, study and therapy of chronic inflammatory joint
XX PT and other diseases, comprises any of many specified genes or derived
XX PT proteins.
XX
XX
XX Claim 1; Page; 12pp; German.
XX
XX
XX This invention describes a novel reagent for diagnosis, molecular
XX definition and therapy of chronic inflammatory joint diseases, and other
XX CC inflammatory disorders, infective or tumour diseases in humans. The
XX CC products of the invention have antinflammatory, cytosratic,
XX antirheptic, antirheumatic and immunosuppressive activity and can be
XX used for gene therapy. The reagent of the invention and any proteins and
XX CC antibodies derived from it, are used (i) for analysing tissue and blood
XX CC samples for medical diagnosis; (ii) for diagnosis and characterisation of
XX CC chronic joint diseases, on the basis of molecular characterisation, and
XX CC determining the etiological pathogenicity principle of as yet
XX CC uncharacterised inflammatory diseases, also monitoring progression and/or
XX CC treatment of disease, and optimisation of therapy and (iii) for
XX CC developing treatments for inflammatory diseases, particularly of joints,
XX CC infections and tumours. ACA64801-ACA64965 represent human polynucleotides
XX CC used in the method of the invention
XX
XX
XX Sequence 2377 BP; 397 A; 862 C; 736 G; 382 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.52e-09 Length: 2377
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.01% Indels: 0
DB: 7 Gaps: 0
US-10-029-345A-109 (1-665) x ACA64956 (1-2377)
QY: 242 VALHISQYsleuAlGlyllleSerArgSerAlaThrllleAlaIleAlaTyrllleWetly 261

```

Db 864 GTCCAGCTGTGGCTGGCATCTCCCGCTCTGCCACCATCCGCACTCCCTCATCATGAG 923

RESULT 33
ID ABX10760 standard; DNA; 2377 BP.
XX
XX ABX10760;
XX
XX 10-MAY-2003 (first entry)
XX
XX Human dual specific phosphatase 8 DNA.
XX
XX Human dual specific phosphatase 8; gene; ds; infection; inflammation;
XX
XX tumour formation; cytostatic; antiinflammatory.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 135..2012
FT CDS /*tag= a
FT /product= "Human dual specific phosphatase 8"
XX
XX US6482644-BI.
XX
XX 19-NOV-2002.
XX
XX 01-AUG-2001; 2001US-00920668.
XX
XX 01-AUG-2001; 2001US-00920668.
XX
XX 01-AUG-2001; 2001US-00920668.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Cowser LM;
XX
XX MPI; 2003-298140/29.
XX
XX P-PSDB; AEG73440.
XX
XX New antisense compound targeted to a nucleic acid encoding human dual
XX
XX specific phosphatase 8, for modulating gene expression and treating
XX
XX diseases associated with expression of the phosphatase in humans.
XX
XX Claim 1; Col 49-54; 36pp; English.
XX
XX The invention relates to a compound targeted to the coding region of a
XX
XX nucleic acid encoding human dual specific phosphatase 8, where the
XX
XX compound specifically hybridises with the region and inhibits the
XX
XX expression of human dual specific phosphatase 8. The compound is useful
XX
XX for inhibiting the expression of human dual specific phosphatase 8 in
XX
XX cells or tissues, and for treating an animal, particularly a human,
XX
XX suspected of having or being prone to a disease or condition associated
XX
XX with expression of dual specific phosphatase 8. The compound is useful
XX
XX for diagnosis, therapeutics and as a research reagent, e.g. to prevent
XX
XX or delay infection, inflammation or tumour formation, and to distinguish
XX
XX between functions of various members of a biological pathway. This
XX
XX sequence represents DNA encoding human dual specific phosphatase 8
XX
XX

Alignment Scores:
Pred. No.: 3.52e-09 Length: 2377
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.01% Indels: 0
DB: 7 Gaps: 0

US-10-029-345A-109 (1-665) x ABX10760 (1-2377)

Qy 242 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 261
Db 864 GTCCAGCTGTGGCTGGCATCTCCCGCTCTGCCACCATCCGCACTCCCTCATCATGAG 923

RESULT 34
ID AAT86757 standard; cDNA; 2415 BP.
XX
XX AAT86757;
XX
XX 15-DEC-1997 (first entry)
XX
XX cDNA of the M3/6 gene.
XX
XX murine; phosphatase; M3/6; dual specificity. Threonine; Tyrosine;
XX
XX inactivate; mitogen activated protein kinase; MAP-K; cdc25 PRP; yeast;
XX
XX trinucleotide repeat; repeat expansion; neurodegenerative; proliferative;
XX
XX diagnosis; tumour; lung; brain; chromosomal deletion; ss.
XX
XX Mus sp.
XX
XX WO9706245-A1.
XX
XX 20-FEB-1997.
XX
XX 05-AUG-1996; 96WO-GB001906.
XX
XX 04-AUG-1995; 95GB-00016059.
XX
XX (MEDI-) MEDICAL RES CONNCTL.
XX
XX Davies KE, Theodosiou A;
XX
XX MPI; 1997-154253/14.
XX
XX Murine mitogen activated protein (MAP) kinase phosphatase M3/6 - is a
XX
XX suspected dual specificity Thr-Tyr phosphatase, useful for diagnosing and
XX
XX treating neuro-degenerative or proliferative diseases e.g. tumours.
XX
XX Claim 1; Fig 1; 51pp; English.
XX
XX This cDNA is said to encode a murine phosphatase designated M3/6 (see
XX
XX also AAT86758). M3/6 is a suspected dual specificity Threonine-Tyrosine
XX
XX phosphatase, capable of inactivating mitogen activated protein (MAP)
XX
XX kinase. The M3/6 protein product shows high homology to the cdc25 PRP of
XX
XX yeast at residues 29-49 and 117-136. The gene also contains a complex
XX
XX tripler distal to the catalytic domain which is translated into the
XX
XX protein. This domain comprises a run of 4 serine residues which in turn
XX
XX is followed by a further run comprising 23 serine residues which is
XX
XX interrupted near the N-terminal section by a single asparagine. This
XX
XX makes the phosphatase gene a candidate for a human disease caused by
XX
XX repeat expansion or mutation. M3/6 is expressed highly in the brain and
XX
XX may have utility in investigating signal transduction mechanisms in brain
XX
XX and muscle. The M3/6 and Hb5 (a human homologue) genes may be
XX
XX responsible, if mutated, for various neurodegenerative or proliferative
XX
XX diseases, and may therefore be used for the diagnosis of such diseases,
XX
XX e.g. tumours, especially lung or brain tumours, associated with deletion
XX
XX of the chromosomal region 11p15.5. The polypeptides can be used to screen
XX
XX for inhibitors to treat these diseases
XX
XX

Alignment Scores:
Pred. No.: 3.57e-09 Length: 2415
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.01% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-109 (1-665) x AAT86757 (1-2415)

Qy 242 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 261
Db 825 GTTCAGTGTGGCTGGCATCTCCCGCTCTGCCACCATCCGCACTCCCTCATCATGAG 884

RESULT 35

AAAT86758
 ID AAT86758 standard; cDNA; 2453 BP.
 AC AAT86758;
 XX
 XX 15-DEC-1997 (first entry)
 DE cDNA of the M3/6 gene.
 XX
 XX murine; phosphatase; M3/6; dual specificity; Threonine; Tyrosine;
 KM inactivate; mitogen activated protein kinase; MAP-K; cdc25 pTP; yeast;
 KM trinucleotide repeat; repeat expansion; neurodegenerative; proliferative;
 KM diagnosis; tumour; lung; brain; chromosomal deletion; ss.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT 98..2089
 FT CDS /*tag= a
 FT /product= "dual-specific_phosphatase"
 XX
 PN MO9706245-A1.
 XX
 PD 20-FEB-1997.
 XX
 PF 05-AUG-1996; 96MO-GB001906.
 XX
 PR 04-AUG-1995; 95GB-00016059.
 XX
 PA (MED1-) MEDICAL RES COUNCIL.
 XX
 PI Davies KE, Theodoros A;
 XX
 DR WPI; 1997-154253/14.
 DR P-PSDB; AAM29150.
 PT Murine mitogen activated protein (MAP) kinase phosphatase M3/6 - is a
 PT suspected dual specificity Thr-Tyr phosphatase, useful for diagnosing and
 PT treating neuro-degenerative or proliferative diseases e.g. tumours.
 XX
 PS Claim 21; Fig 2; 51pp; English.
 XX
 CC This cDNA encodes a murine phosphatase designated M3/6 (see also
 CC AAT86757). M3/6 is a suspected dual specificity Threonine-Tyrosine
 CC phosphatase, capable of inactivating mitogen activated protein (MAP)
 CC kinase. The M3/6 protein product shows high homology to the cdc25 pTP of
 CC yeast at residues 29-49 and 117-136. The gene also contains a complex
 CC triplet distal to the catalytic domain which is translated into the
 CC protein. This domain comprises a run of 4 serine residues which in turn
 CC is followed by a further run comprising 23 serine residues which is
 CC interrupted near the N-terminal section by a single asparagine. This
 CC makes the phosphatase gene a candidate for a human disease caused by
 CC repeat expansion or mutation. M3/6 is expressed highly in the brain and
 CC may have utility in investigating signal transduction mechanisms in brain
 CC and muscle. The M3/6 and Hb5 (a human homologue) genes may be
 CC responsible, if mutated, for various neurodegenerative or proliferative
 CC diseases, and may therefore be used for the diagnosis of such diseases,
 CC e.g. tumours, especially lung or brain tumours, associated with deletion
 CC of the chromosomal region 11p15.5. The polypeptides can be used to screen
 CC for inhibitors to treat these diseases
 XX
 SO Sequence 2453 BP; 493 A; 784 C; 705 G; 471 T; 0 U; 0 Other;

Alignment Scores:

Prod. No.: 3.62e-09 Length: 2453
 Score: 20.00 Matches: 20
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.01% Indels: 0
 DB: 2 Gaps: 0

US-10-029-345A-109 (1-665) X AAT86758 (1-2453)

OY 242 VALHICVLEuAAGLIYIleSeraGSeRaLaThrIleAlaIleAlaTyrlIleMeLyS 261
 DB 827 GTTCACTGCTGCGTCGATCTCTGCTGCGACCATTCGCCATCCGATCATCATGAA 886
 RESULT 36
 ID AAD41236
 XX AAD41236 standard; cDNA; 2453 BP.
 AC AAD41236;
 XX
 XX 30-OCT-2002 (first entry)
 DE Murine neuronal tyrosine/threonine phosphatase (NTPP1) cDNA.
 XX
 XX Murine; targeting construct; neuronal tyrosine/threonine phosphatase;
 KM NTPP1 gene; transgenic mouse; pharmacological therapy; genetic disease;
 KM neurological; neuropsychological; psychotic illness; ss.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 1..142
 FT /*tag= a
 FT /note= "Sequence flanking Neo insert in targeting
 FT construct"
 FT misc_feature 143..304
 FT /*tag= b
 FT /note= "Deleted in targeting construct"
 FT misc_feature 305..328
 FT /*tag= c
 FT /note= "Sequence flanking Neo insert in targeting
 FT construct"
 XX
 PN WO200245497-A2.
 XX
 PD 13-JUN-2002.
 XX
 PP 05-DEC-2001; 2001MO-US046868.
 XX
 PR 06-DEC-2000; 2000US-0251802P.
 PR 04-DEC-2001; 2001US-00005858.
 XX
 PA (DELTA-) DELTAGEN INC.
 XX
 PI Allen KD;
 XX
 DR WPI; 2002-537527/57.
 PT Novel targeting construct for producing a transgenic mouse useful as a
 PT disease model, has nucleotide sequences homologous to a neuronal
 PT tyrosine/threonine phosphatase gene and a selectable marker.
 XX
 PS Example 1; Fig 2A; 48pp; English.
 XX
 CC The invention relates to a targeting construct comprising a first
 CC polynucleotide sequence homologous to at least a first portion of
 CC neuronal tyrosine/threonine phosphatase (NTPP1) gene, a second
 CC polynucleotide sequence homologous to at least a second portion of the
 CC (NTPP1) gene and a selectable marker. The invention is useful for
 CC producing a transgenic mouse comprising a disruption in a NTPP1 gene. The
 CC invention is useful for identifying an agent that ameliorates a phenotype
 CC associated with a disruption in a NTPP1 gene. The invention is useful for
 CC relating a unique animal model for testing and developing new treatments
 CC relating to the behavioural phenotypes, which can be used to test the
 CC efficacy of proposed genetic and pharmacological therapies for human
 CC genetic disease, such as neurological, neuropsychological, or psychotic
 CC illnesses. The invention is utilized as models for diseases, disorders,
 CC or conditions associated with phenotypes relating to a disruption in a
 CC NTPP1 gene, to identify drugs, pharmaceuticals, drugs, therapies and
 CC interventions that are effective in treating disease. The present
 CC sequence is murine NTPP1 cDNA
 XX
 SO Sequence 2453 BP; 493 A; 784 C; 705 G; 471 T; 0 U; 0 Other;

Alignment Scores:		Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
US-10-029-345A-109 (1-665) x AAD41236 (1-2453)		3.62e-09	20.00	100.00%	100.00%	3.01%	6
QY	242 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetIys 261						
Db	827 GTTCACTGCTGGCTGGCATCTCTCGCTTGCACCAATTCGATCGGTCATCATGAA 886						
RESULT 37							
AAS31013							
ID	AAS31013 standard; cDNA; 2476 BP.						
XX							
AC	AAS31013;						
XX							
DT	04-DEC-2001 (first entry)						
XX							
DE	Human diagnostic and therapeutic polynucleotide (DITHP) #28.						
XX							
KW	Human, receptor; diagnostic; therapeutic; gene therapy; vaccine; cell proliferative disorder; Crohn's disease; lymphoma; leukaemia; acquired immune deficiency syndrome; AIDS; autoimmune disorder; respiratory disorder; ss.						
XX							
OS	Homo sapiens.						
XX							
PN	WO200162927-A2.						
XX							
PD	30-AUG-2001.						
XX							
PF	21-FEB-2001; 2001WO-US006059.						
XX							
PR	24-FEB-2000; 2000US-0184693P.						
PR	24-FEB-2000; 2000US-0184697P.						
PR	24-FEB-2000; 2000US-0184698P.						
PR	24-FEB-2000; 2000US-0184768P.						
PR	24-FEB-2000; 2000US-0184769P.						
PR	24-FEB-2000; 2000US-0184770P.						
PR	24-FEB-2000; 2000US-0184771P.						
PR	24-FEB-2000; 2000US-0184772P.						
PR	24-FEB-2000; 2000US-0184773P.						
PR	24-FEB-2000; 2000US-0184774P.						
PR	24-FEB-2000; 2000US-0184776P.						
PR	24-FEB-2000; 2000US-0184777P.						
PR	24-FEB-2000; 2000US-0184797P.						
PR	24-FEB-2000; 2000US-0184813P.						
PR	24-FEB-2000; 2000US-0184837P.						
PR	24-FEB-2000; 2000US-0184841P.						
PR	24-FEB-2000; 2000US-0185213P.						
PR	24-FEB-2000; 2000US-0185216P.						
PR	12-MAY-2000; 2000US-0203785P.						
PR	15-MAY-2000; 2000US-0203785P.						
PR	16-MAY-2000; 2000US-0204226P.						
PR	16-MAY-2000; 2000US-0204226P.						
PR	16-MAY-2000; 2000US-0204821P.						
PR	16-MAY-2000; 2000US-0204908P.						
PR	16-MAY-2000; 2000US-0204821P.						
PR	16-MAY-2000; 2000US-0205233P.						
PR	17-MAY-2000; 2000US-0204815P.						
PR	17-MAY-2000; 2000US-0204863P.						
PR	17-MAY-2000; 2000US-0205221P.						
PR	17-MAY-2000; 2000US-0205285P.						
PR	17-MAY-2000; 2000US-0205286P.						
PR	17-MAY-2000; 2000US-0205287P.						
PR	17-MAY-2000; 2000US-0205323P.						
XX							
XX	17-MAY-2000; 2000US-0205324P.						

PI Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC,
 PI Chen A, D'sa SA, Amshey S, Dahl CR, Dam TC, Daniels SE, Dautour GB;
 PI Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL, Liu TF,
 PI Rosberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A;
 PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W,
 PI Cohen HU, Hodgson DM, Lincoln SE, Jackson S;
 DR MPI: 2001-502867/55.
 DR P-PDB: AAU19442.
 XX
 PT polynucleotides encoding diagnostic and therapeutic proteins, e.g.
 PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.
 XX
 PS
 XX
 XX Claim 1; Page 310; 522pp; English.
 CC The invention relates to polynucleotides (I) encoding diagnostic and
 CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and
 CC proteins involved in growth and development and receptors. (I) and (II)
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate DITHP expression. For example, (I) and (II)
 CC may be used to treat disorders associated with decreased polypeptide
 CC expression by rectifying mutations or deletions in a patient's genome,
 CC that affect the activity of the DITHPs. By expressing inactive proteins
 CC or supplementing the patient's own production of them. (I) and (II) may
 CC be used to treat diseases, for example, cell proliferative disorder,
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,
 CC leukemia, autoimmune disorders, and respiratory disorders. Additionally
 CC (I) may be used to produce the DITHPs, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. (I) and
 CC its complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. (II) may also be used as antigens in the production of
 CC antibodies against DITHPs and in assays to identify modulators of DITHP
 CC expression and activity. The anti-DITHP antibodies and antagonists may
 CC also be used to down regulate expression and activity. The anti-DITHP
 CC antibodies may also be used as diagnostic agents for detecting the
 CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA)). AAS30986-AAS31196 represent human diagnostic and therapeutic
 CC (DITHP) polynucleotides of the invention
 XX
 XX
 SQ Sequence 2476 BP; 432 A; 884 C; 762 G; 398 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3.66e-09 Length: 2476
 Score: 20.00 Matches: 20
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 3.01% Indels: 0
 Ds: 4 Gaps: 0
 US-10-029-345A-109 (1-665) x AAS31013 (1-4476)
 Oy 242 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 261
 Db 933 GTCACCTGCTGGCTGCATCTCCGCTCTGCACCATCGCCATCGCTCATCATGAG 9922
 RESULT 38
 ABL13421
 XX ID ABL13421 standard; cDNA; 2369 BP.
 XX ABL13421,
 AC
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 34745.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 MN MO200171042-A2.

XX 27-SEP-2001.
 PD ABLJ3420 standard; cDNA; 18413 BP.
 PF 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR P-PSDB; ABB69318.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Claim 1; SEQ ID NO 34745; 21bp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 SQ Sequence 2369 BP; 716 A; 591 C; 557 G; 505 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.0285 Length: 2369
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.95% Indels: 0
 DB: Gaps: 0
 US-10-029-345A-109 (1-665) x ABLJ3421 (1-2369)
 QY 246 AAGAGTCTTGGGATTTCCGCGAGCGCCACCATTCGCTTAC 258
 ID ABLJ3420 standard; cDNA; 18413 BP.
 DB 1125 GCGGGATTTCCGCGAGCGCCACCATTCGCTTAC 1163
 RESULT 39
 ABLJ3420
 ID ABLJ3420 standard; cDNA; 18413 BP.
 XX ABLJ3420;
 AC 26-MAR-2002 (first entry)
 DT *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 34742.
 DE *Drosophila*; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 KM *Drosophila melanogaster*.
 OS WO200171042-A2.
 PN 27-SEP-2001.
 PD 23-MAR-2001; 2001WO-US009231.
 PF 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR P-PSDB; ABB69317.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Claim 1; SEQ ID NO 34742; 21bp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 SQ Sequence 18413 BP; 5536 A; 3801 C; 3766 G; 5310 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.202 Length: 18413
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.95% Indels: 0
 DB: Gaps: 0
 US-10-029-345A-109 (1-665) x ABLJ3420 (1-18413)
 QY 246 AAGAGTCTTGGGATTTCCGCGAGCGCCACCATTCGCTTAC 258
 ID ABLJ3420 standard; cDNA; 18413 BP.
 DB 16169 GCGGGATTTCCGCGAGCGCCACCATTCGCTTAC 16207
 RESULT 40
 ABLJ3420
 ID ABLJ3420 standard; cDNA; 18413 BP.
 XX ABLJ3420;
 AC 23-JAN-2003 (first entry)
 DT Human dual specificity protein phosphatase-7-like protein cDNA.
 DE Dual-specificity protein phosphatase 7-like protein; DSGP7; human;
 KW protein phosphatase; enzyme; cancer; asthma; obesity; diabetes;
 KW cardiovascular disorder; chronic obstructive pulmonary disorder;
 KW antitumor; antiparkinsonian; nootropic; neuroprotective; cardiac;
 KW hypotensive; antiarrhythmic; analgesic; antiallergic;
 KW antiinflammatory; anorectic; antidiabetic; tranquilizer; antimanic;
 KW antidepressant; dermatological; gene therapy; chromosome 3p21.1; gene;
 KW ss.
 KM Homo sapiens.
 OS 179.1435
 PN 03-OCT-2002.
 PD 20-MAR-2002; 2002WO-EP003091.
 PF 20-MAR-2002; 2002WO-EP003091.

XX 23-MAR-2001; 2001US-0277954P.
 PR 28-JAN-2002; 2002US-0351377P.
 XX
 XX (FARB) BAYER AG.
 PA
 PI Zhu Z;
 XX
 XX MPI: 2003-029935/02.
 DR P-ESDB; ABB55026.
 XX
 XX
 PT New dual specificity protein phosphatase (DSPP)7-like protein polypeptide
 PR and polynucleotide, useful for regulating DSPP7-like protein activity for
 PT preventing, treating or ameliorating cancer or cardiovascular disorders.
 PS
 PS Claim 19; Fig 4; 132pp; English.
 XX
 XX
 CC The present sequence is that of a cDNA clone, the coding region of which
 CC is also claimed, encoding a novel human dual specificity protein
 CC phosphatase 7 (DSPP7)-like protein. The sequence is located on chromosome
 CC 3p21.1. The invention provides human DSPP7-like polypeptides and
 CC polynucleotides, and methods of using them for identifying reagents which
 CC regulate DSPP7-like protein or which bind to DSPP7-like gene products.
 CC These can be used in preventing, ameliorating or correcting dysfunctions
 CC or diseases associated with DSPP7-like protein dysfunction such as
 CC cancer, central nervous system disorders (e.g. mood disorders, anxiety,
 CC Parkinson's disease, Alzheimer's disease), asthma (including other
 CC allergic diseases such as allergic rhinitis or atopic dermatitis),
 CC cardiovascular disorders (e.g. myocardial infarction, hypertension,
 CC arrhythmias or angina pectoris), chronic obstructive pulmonary disease,
 CC obesity, or diabetes (all claimed). They can also be used to treat pain
 CC associated with the above disorders. The polynucleotides can also be used
 CC in diagnostic assays or in genetic testing, as probes or primers, in the
 CC production of DSPP7-like recombinant polypeptides, in gene therapy,
 CC antisense oligonucleotides and ribozymes are used in a claimed method of
 CC reducing human DSPP7-like protein activity
 XX
 XX
 SQ Sequence 1435 BP; 252 A; 517 C; 452 G; 207 T; 0 U; 7 Other;
 Alignment Scores:
 Pred. No.: 0.171 Length: 1435
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 1.80% Indels: 0
 Db: 7 Gaps: 0
 US-10-029-345A-109 (1-665) x ABV74143 (1-1435)
 QY 240 ValLeuValHisCysLeuNlaIylIseerYgsr 251
 Db 1157 GTCTGTGTCATGCTGCTGGCAAGGCAATCAGCCGCTCA 1192
 RESULT 41
 ID AAD57367 standard; cDNA; 1471 BP.
 AC AAD57367;
 XX
 XX AAD57367;
 DT 06-NOV-2003 (first entry)
 XX
 XX
 DE Human kinase and phosphatase (KPP-40) cDNA.
 XX
 XX Human; kinase; phosphatase; KPP; cell proliferative disorder; hepatitis;
 KM atherosclerosis; cirrhosis; hemoglobinuria; polycythemia vera; cancer;
 KM psoriasis; thrombocytopaenia; developmental disorder; Reiter's syndrome;
 KM renal tubular acidosis; anaemia; mental retardation; Alzheimer's disease;
 KM neurological disorder; Parkinson's disease; epilepsy; glomerulonephritis;
 KM autoimmune disorder; inflammatory disorder; AIDS; Goodpasture's syndrome;
 KM acquired immune deficiency syndrome; immunosuppressive; Crohn's disease;
 KM noctropic; transgenic; dermatitis; multiple sclerosis; diabetes mellitus;
 KM allergy; gout; Grave's disease; Hashimoto's thyroiditis; bowel syndrome;
 KM osteoporosis; rheumatoid arthritis; Sjogren's syndrome; ophthalmological;

XX	Gene therapy; aschma; anticonvulsant; uropathic; pancreatitis; gene; ss.
OS	Homo sapiens.
FH	Key
FT	Location/Qualifiers
FT	CDS
XX	170..690
XX	/tag= a
XX	/product= "Human KPP"
PN	WO2003050084-A2.
XX	
PD	19-JUN-2003.
XX	
PF	06-DEC-2002; 2002MO-US039126.
XX	
PR	07-DEC-2001; 2001US-0340235P.
PR	19-DEC-2001; 2001US-0343007P.
PR	21-DEC-2001; 2001US-0343546P.
PR	04-FEB-2002; 2002US-0354388P.
PR	15-FEB-2002; 2002US-0357675P.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
PI	
PI	Kable AE, Chien D, Wilson AD, Swarnakar A, Gorvad AE;
PI	Hafalia AA, Emerling BM, Ramkumar J, Jin P, Griffin JA, Marquis JP,
PI	Baugh MR, Chawla NK, Lehr-Mason PM, Khare R, Lee S, Hawkins PR;
PI	Becha SD, Lee SY, Sprague WW, Zebzarjadian Y;
DR	WPI: 2003-532894/50.
DR	P-PSDB; AAE37995.
XX	
PT	New human kinases and phosphatases and polynucleotides, useful for
PT	diagnosing, treating or preventing autoimmune or inflammatory disorders
PT	(e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
PT	cancer or hepatitis.
XX	
PS	Claim 5; Page 281; 282pp; English.
XX	
CC	The invention relates to an isolated polypeptide, which is a human kinase
CC	and phosphatase (KPP). KPP agonists and antagonists are useful for
CC	diagnosing, treating or preventing disorders associated with aberrant
CC	expression of KPP, particularly cell proliferative disorders (e.g.
CC	arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal
CC	nocturnal haemoglobinuria, polycythaemia vera, psoriasis, primary
CC	thrombocytopoenia or cancer), developmental disorders (eg. renal tubular
CC	acidosis, anaemia or mental retardation), neurological disorders (e.g.
CC	Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/
CC	inflammatory disorders (e.g. AIDS, acquired immune deficiency syndrome,
CC	allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's
CC	disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome,
CC	gout, Grave's disease, Hashimoto's thyroiditis, irritable bowel syndrome,
CC	multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
CC	syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
CC	bacterial, fungal, parasitic, protozoan or helminthic infections. The KPP
CC	is useful in assessing the effects of exogenous compounds on the
CC	expression of nucleic acids and kinases and phosphatases. KPP gene is
CC	useful in gene therapy and for creating transgenic animals to model human
CC	disease. The present sequence is human KPP CDNA
XX	
SQ	Sequence 1471 BP; 295 A; 479 C; 384 G; 313 T; 0 U; 0 Other;
XX	
Alignment Scores:	
Pred. No.:	0.175
Score:	12.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	1.80%
DB:	8
	Gaps: 0
US-10-029-345A-109 (1-665) x AAD57367 (1-1471)	
240 ValLeuValHisCysLeuAlaGlyIleSerArgSer 251	

Dn			618 GTCCTGTGCATGCGCAGCATCACGCCGTCA	653
	RESULT 42			
Xx	AAD57359	ID	AAD57359 standard; cDNA; 1472 BP.	
Xx		AC		
Xx	AAD57359;			
Dt				
Dt	06-NOV-2003	(first entry)		
Xx				
Dd				
Xx				
Xx				
Kw			Human kinase and phosphatase (KPP-32) CDNA.	
Km			Human; kinase; phosphatase; KPP; cell proliferative disorder; hepatitis; atherosclerosis; cirrhosis; haemoglobinuria; polycythaemia vera; cancer; psoriasis; thrombocytopenia; developmental disorder; Reiter's syndrome; renal tubular acidosis; anaemia; mental retardation; Alzheimer's disease; neurological disorder; Parkinson's disease; epilepsy; glomerulonephritis; autoimmune disorder; inflammatory disorder; AIDS; Goodpasture's syndrome; acquired immune deficiency syndrome; immunosuppressive; Crohn's disease; nocrotic; traumatic; dermatitis; multiple sclerosis; diabetes mellitus; allergy; gout; Grave's disease; Hashimoto's thyroiditis; bowel syndrome; osteoporosis; rheumatoid arthritis; Sjogren's syndrome; ophthalmological; gene therapy; asthma; anticonvulsant; uropathic; pancreatitis; gene; ss.	
Xk	Homo sapiens.			
Oo				
Fh	Key	Location/Qualifiers		
Ff	CDS	17..538 /*tag= a		
Pt		/product= "Human KPP"		
Xc				
Nn	M02003050084-A2.			
Xd				
Pd	19-JUN-2003.			
Pf	06-DEC-2002; 2002MOC-US039126.			
Xg				
Ee	07-DEC-2001; 2001US-0340235P. 19-DEC-2001; 2001US-0343007P. 21-DEC-2001; 2001US-0343546P. 04-FEB-2002; 2002US-0354388P. 15-FEB-2002; 2002US-0357675P.			
Rr				
Pa	(INCY-) INCYTE GENOMICS INC.			
Xl				
Pi	Kable AE, Chien D, Wilson AD, Swarnakar A, Gorvyad AB; Hafeela AJA, Emerling BM, Ramnumar J, Jin P, Griffin JA, Marquis JP, Baughn MR, Chavla NK, Lehr-Mason PM, Khare R, Lee S, Hawkins PR; Becha SD, Lee SY, Sprague WM, Zebardjadian Y;			
Xx				
Dz	WPI; 2003-532894/50.			
Dr	P-PsDB; AAEE37987.			
Xx				
Pt	New human kinases and phosphatases and polymucleotides, useful for diagnosing, creating or preventing autoimmune or inflammatory disorders (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis, cancer or hepatitis.			
Xt				
Pt				
Xs	Claim 5; Page 275; 282pp; English.			
The invention relates to an isolated polypeptide, which is a human kinasees and phosphatase (KPP). KPP agonists and antagonists are useful for diagnosing, treating or preventing disorders associated with aberrant expression of KPP, particularly cell proliferative disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia vera, psoriasis, primary thrombocytopoenia or cancer), developmential disorders (eg. renal tubular acidosis, anaemia or mental retardation), neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/ inflammatory disorders (e.g. AIDS; acquired immune deficiency syndrome, allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus, glomerulonephritits, Goodpasture's syndrome, gout, Grave's disease, Hashimoto's thyroiditits, irritable bowel syndrome,				

CC	multiple sclerosis, osteoarthritis, osteoporosis, pancreaticitis, Reiter's
CC	syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
CC	bacterial, fungal, parasitic, protozoan or helminthic infections. The KPP
CC	is useful in assessing the effects of exogenous compounds on the
CC	expression of nucleic acids and kinases and phosphatases. KPP gene is
CC	useful in gene therapy and for creating transgenic animals to model human
CC	disease. The present sequence is human KPP cDNA
XX	
XX	Sequence 1472 BP, 295 A; 479 C; 385 G; 313 T; 0 U; 0 Other;
US-10-029-345A-109 (1-665) x AAD57359 (1-1472)	
Qy	240 ValLeuValHisCysLeuAlaGlyIleSerArgSer 251
Db	618 GTCCTGTCACACTGCCTGGCAGCATCAGCCGCTCA 653
RESULT 43	
ADCG9150	standard; cDNA; 1547 BP.
XX	
XX	ADCG9150;
XX	
DT	01-JAN-2004 (first entry)
XX	
DE	Human KPP cDNA - SEQ ID 103.
XX	
XX	anti-HIV; antiallergic; antiinflammatory; antianaemic; antiparkinsonian;
KM	neurotropic; anticonvulsant; antiarteriosclerotic; antischismatic;
KM	immunosuppressive; antichyroid; cyostatic; hepatotropic; dermatological;
KM	antidiabetic; nephrotropic; antigout; thyromimetic; neuroprotective;
KM	osteoplastic; antiarthritic; antiparasitic; antihelminthic; antiparasitic;
KM	uropathic; ophthalmological; antihemetic; haemostatic; antibacterial;
KM	virucide; protozoicide; fungicide; kinase; phosphatase; KPP;
KM	cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis;
KM	cancer; developmental; mental retardation; neurological;
KM	Alzheimer's disease; Parkinson's; autoimmune; inflammatory; Crohn's;
KM	diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;
KM	helminthic infection; transgenic; gene therapy; human; ss; gene.
OS	
XX	Homo sapiens.
XX	
XX	MO2003033680-A2.
XX	
PN	
PD	24-APR-2003.
XX	
PF	17-OCT-2002; 2002WO-US033723.
XX	
PR	19-OCT-2001; 2001US-0345474P.
PR	02-NOV-2001; 2001US-0343910P.
PR	13-NOV-2001; 2001US-0330989P.
PR	16-NOV-2001; 2001US-0332424P.
PR	30-NOV-2001; 2001US-0334288P.
XX	
XX	
XX	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Bandman O., Baughn MR., Becha SD., Borowsky ML., Duggan BM.
PI	Emerling BM., Forsythe LJ., Geddis AR., Gotzard AE., Griffin JA.
PI	Gururajan R., Hatala AD., Khan FA., Lal FG., Lee EA., Lee ST.
PI	Lindquist EA., Lu DM., Lu Y., Marquis JP., Nguyen DB., Arvizu CS.
PI	Rinkumar J., Reardon SA., Richardson TW., Swarnakar A., Tang YT.
PI	Thornton MB., Tran UK., Chawla NK., Warren BA., Yang J., Yue H.
XX	Zebrafishian Y.
XX	
XX	WPI, 2003-403214/38.
DR	P-PSDB; ADCG9098.

DE		Human colon cancer antigen nucleotide sequence SFG ID NO:69.
XX	KM	Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX	KM	identification; cytostatic; cardioactive; neuroprotective; vulnery;
XX	KM	immunomodulatory; muscular; gynaecological; gastrointestinal;
XX	KW	nephrotropic; antiinfective; antibacterial; gene therapy; wound;
XX	KV	neural disorder; immune system disorders; muscular disorder;
XX	KM	reproductive disorder; gastrointestinal disorder; renal disorder;
XX	KX	infectious disease; cardiovascular disorder; ss.
XX	OS	Homo sapiens.
XX	PN	WO200055351-A1.
XX	PD	21-SEP-2000.
XX	PF	08-MAR-2000; 2000WO-US005883.
XX	PR	12-MAR-1999; 99US-0124270P.
XX	PA	(HUMA-) HUMAN GENOME SCI INC.
XX	P1	Rosen CA, Ruben SM;
XX	DR	WPI: 2000-587534/55.
XX	P7	P-PSTDB; AAB53302.
XX	PT	Colo cancer associated gene sequences, referred to as colo cancer
XX	PS	disorders such as colo cancer.
XX	PS	Claim 1; Page 517-518; 2104pp; English.
XX	CC	AAC97991 to AAC98763 encode the human colo cancer associated proteins,
XX	CC	called human colo cancer antigens, given in AAB53234 to AAB54006. The
XX	CC	human colo cancer antigen can have cytostatic, cardioactive, muscular,
XX	CC	neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
XX	CC	vulnery, nephrotropic, antiinfective and antibacterial activities, and
XX	CC	can be used in gene therapy. The colo cancer antigen polynucleotides,
XX	CC	proteins and antibodies to the proteins are useful for the prevention,
XX	CC	treatment and diagnosis of colo disorders, such as colo cancer. The
XX	CC	polynucleotides may be used in diagnostics and research, such as for
XX	CC	chromosome identification, and as hybridisation probes. The proteins may
XX	CC	also be used to prevent diseases such as neural disorders, immune system
XX	CC	disorders, muscular disorders, reproductive disorders, gastrointestinal
XX	CC	disorders, wound, renal disorders, infectious diseases, and
XX	CC	cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
XX	CC	sequences used in the exemplification of the present invention
XX	SQ	Sequence 1848 BP, 567 A, 359 C, 378 G, 539 T, 0 U, 5 Other;
XX		
XX		Alignment Scores:
XX	Pred. No.:	0.218 Length: 1848
XX	Score:	12.00 Matches: 12
XX	Percent Similarity:	100.00% Conservative: 0
XX	Best Local Similarity:	100.00% Mismatches: 0
XX	Query Match:	1.80% Indels: 0
XX		Gaps: 0
XX		
XX	US-10-029-345A-109 (1-665) x AAC98059 (1-1848)	
XX	Oy	240 ValIeuValHisCysLeuAlaGlyIleSerArgSer 251
XX	Dd	290 GTCTTGGAATTCCTTGGCGGATTAAGCGGTCA 325
XX		
XX	RESULT 46	
XX	ABT42339	
XX	ID	ABT42339 standard; DNA; 2104 BP.
XX	AC	ABT42339;
XX	XX	
XX	DT	26-JUN-2003 (first entry)
XX	XX	

DE	Toxicity modelling related rat gene SEQ ID No 241.
XX	
KM	Toxic effect; gene expression profile; renal toxicity; toxicity marker;
KW	database; drug screening; toxicity assay; rat; ds.
XX	
OS	Rattus norvegicus .
XX	
PN	WO200295000-A2.
XX	
PD	28-NOV-2002.
XX	
PF	22-MAY-2002; 2002WO-US016173.
XX	
PR	22-MAY-2001; 2001US-0292335P.
PR	13-JUN-2001; 2001US-0297523P.
PR	19-JUN-2001; 2001US-0298925P.
PR	10-JUL-2001; 2001US-0303807P.
PR	10-JUL-2001; 2001US-0303808P.
PR	10-JUL-2001; 2001US-0303810P.
PR	28-AUG-2001; 2001US-0315047P.
PR	27-SEP-2001; 2001US-0324928P.
PR	22-OCT-2001; 2001US-0330462P.
PR	01-NOV-2001; 2001US-0330867P.
PR	21-NOV-2001; 2001US-0331805P.
PR	06-DEC-2001; 2001US-0336144P.
PR	19-DEC-2001; 2001US-0340873P.
PR	21-FEB-2002; 2002US-0357842P.
PR	21-FEB-2002; 2002US-0357843P.
PR	21-FEB-2002; 2002US-0357844P.
PR	15-MAR-2002; 2002US-0364134P.
PR	08-APR-2002; 2002US-0370144P.
PR	08-APR-2002; 2002US-0370206P.
PR	08-APR-2002; 2002US-0370247P.
PR	17-APR-2002; 2002US-0372794P.
PR	21-APR-2002; 2002US-0371679P.
XX	
PA	(GENE-) GENE LOGIC INC.
XX	
PI	Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
DR	WPI, 2003-148464/14.
XX	
PT	Predicting at least one toxic effect of a compound, useful for toxicity
PT	modeling, comprises preparing a gene expression profile of a tissue or
PT	cell sample exposed to the compound, and comparing the gene expression
PT	profile to a database.
XX	
PS	Example 4; Page; 446pp; English.
XX	
CC	The invention relates to a novel method of predicting at least one toxic
CC	effect of a compound. The method comprises a gene expression profile of a
CC	tissue or cell sample exposed to the compound, and comparing the gene
CC	expression profile to a database comprising at least part of the data or
CC	information given in the specification. The methods are useful for
CC	predicting at least one toxic effect of a compound, predicting the renal
CC	progression of a toxic effect of a compound, predicting the renal
CC	toxicity of a compound, or identifying toxicity markers in tissues or
CC	cells exposed to known renal toxin. The genes are useful as toxicity
CC	markers in drug screening and toxicity assays, in monitoring disease or
CC	physiological states, or disease progression. This polynucleotide
CC	represents a rat DNA sequence relating to the toxic effect database
CC	described in the specification. NOTE: The sequence data for this patent
CC	did not form part of the printed specification, but was obtained in
CC	electronic format directly from the World Intellectual Property
CC	Organization
XX	
SQ	. Sequence 2104 BP; 463 A; 559 C; 558 G; 524 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	0.247 Length: 2104
Score:	12.00 Matches: 12
Percent Similarity:	100.00% Conservatve: 0
Best Local Similarity:	100.00% Mismatches: 0

Query Match: 1.80% Indels: 0
DB: 7 Gaps: 0
US-10-029-345A-109 (1-665) x ABT2339 (1-2104)
QY 240 ValLeuValHISCySLeuA1aG1Y11eSerARgSer 251
Db 1225 GTCTGTGTCATTCCTTGGCGGCATCAGCCGCTCC 1260
RESULT 47
ADBS3531
ID ADBS3531 standard; DNA; 2104 BP.
XX
AC ADBS3531;
XX
DT 04-DEC-2003 (first entry)
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:4073.
XX
XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KM toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; de.
XX
OS Rattus norvegicus.
XX
PN WO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
XX 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
DR WPI; 2003-731472/69.
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of tox mean and non-tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 4073; 874pp; English.
XX
XX The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the tox mean and non-tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state

CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 2104 BP; 463 A; 559 C; 558 G; 524 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0.247 Length: 2104
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 9 Gaps: 0
US-10-029-345A-109 (1-665) x ADBS3531 (1-2104)
QY 240 ValLeuValHISCySLeuA1aG1Y11eSerARgSer 251
Db 1225 GTCTGTGTCATTCCTTGGCGGCATCAGCCGCTCC 1260
RESULT 48
ACA56537
ID ACA56537 standard; cDNA; 2109 BP.
XX
AC ACA56537;
XX
DT 06-JUN-2003 (first entry)
DE Human signalling pathway polynucleotide probe SEQ ID NO 1135.
XX
XX Human; probe; ss; array element; Parkinson's disease;
KM signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
OS Homo sapiens.
XX
PN US6500938-B1.
XX
PD 31-DEC-2002.
XX
PF 30-JAN-1998; 98US-00016434.
XX
PR 30-JAN-1998; 98US-00016434.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Au-Young J, Seilhamer JJ;
XX
DR WPI; 2003-352189/33.
XX
PT Combination of polynucleotide probes, useful as array elements in a
PT microarray for monitoring the expression of a number of target
PT polynucleotides.
XX
PS Claim 1; SEQ ID NO 1135; 65pp; English.
XX
XX The invention relates to a combination which, comprises a number of
CC polynucleotide probes comprising a sequence selected from one of the 1490
CC sequences mentioned in the specification. The combination is useful as an
CC array element in a microarray for monitoring the expression of a number
CC of target polynucleotides. The microarray is particularly useful in the
CC diagnosis and treatment of cancer and immunopathology and neuropathology.
CC The microarray is useful in diagnostics and treatment regimens, drug
CC discovery and development, toxicological and carcinogenicity studies,
CC forensics and pharmacogenomics. The microarray is also useful for
CC monitoring progression of diseases and for developing sophisticated
CC profiles for the effects of currently available therapeutic drugs. The
CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC and genomic fragments and in research and diagnostic applications. The
CC array can detect changes in expression in a large number of genes coding
CC for different signalling pathway populations which can be used to diagnose

CC	various diseases including cancer e.g. adenocarcinoma and leukemia,
CC	immunopathies e.g. AIDS and ashrma, neuropathies e.g. Alzheimer's disease
CC	and Parkinson's disease. The present sequence represents a polynucleotide
CC	probe of the invention. Note: The sequence data for this patent did not
CC	form part of the printed specification but was obtained in electronic
CC	format directly from USPTO at
CC	seqdata.uspto.gov/sequence.html?docID=06500938B1
XX	
SO	Sequence 2109 BP; 484 A; 555 C; 561 G; 509 T; 0 U; 0 Other;
	Alignment Scores:
	Pred. No.: 0.248 Length: 2109
	Score: 12.00 Matches: 12
	Percent Similarity: 100.00% Conservativeness: 0
	Best Local Similarity: 100.00% Mismatches: 0
	Query Match: 1.80% Indels: 0
	DB: 7 Gaps: 0
US-10-029-345A-109	(1-665) x ACA56537 (1-2109)
Oy	240 ValLeuValHhCysLeuAlaGlyIleSerArgSer 251
Db	1216 GTCTTGTAATTGCTGCTGGCTGGCATTTAGCGCTCA 1251
RESULT 49	
ABK83457	
ID	ABK83457 standard; CDNA; 2390 BP.
XX	
AC	ABK83457;
XX	
DT	14-AUG-2002 (first entry)
DE	
XX	Human CDNA differentially expressed in granulocytic cells #28.
KW	Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW	vital infection; parasitic infection; protozoal infection;
KW	fungal infection; sterile inflammatory disease; psoriasis;
KW	rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW	cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW	adult respiratory distress syndrome; inflammatory bowel disease;
KW	Crohn's disease; ulcerative colitis; periodontal disease;
KW	granulocyte activation; chronic inflammation; allergy.
XX	
OS	Homo sapiens.
XX	
PN	WO200228999-A2.
XX	
PD	11-APR-2002.
XX	
PF	03-OCT-2001; 2001WO-US030821.
XX	
PR	03-OCT-2000; 2000US-0237189P.
XX	
PA	(GENE-) GENE LOGIC INC.
XX	
P1	Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX	
DR	WPI; 2002-435328/46.
XX	
PT	Detecting granulocyte activation by detecting differential expression of
PT	genes associated with granulocyte activation, which serves as diagnostic
PT	markers that is useful for monitoring disease states and drug toxicity.
XX	
PS	Claim 1; SEQ ID NO 28; 114pp; English.
XX	
CC	The invention relates to detecting (M1) granulocyte (GC) activation
CC	(GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC	DNA chip analysis as given in the specification, and comparing the
CC	expression level to an expression level in an unactivated GC, where
CC	differential expression of Gs is indicative of GCA. Also included are
CC	modulating (M2) GA by contacting GC with an agent that alters the
CC	expression of at least one gene in Gs; (2) screening (M3) for an agent
CC	capable of modulating GCA or an inflammation (especially chronic) in a

```

CC tissue,an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from G6, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from G6 in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC
SQ Sequence 2390 BP: 565 A; 602 C; 616 G; 607 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.279 Length: 2390
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
De: 6 Gaps: 0

US-10-0229-345A-109 (1-665) x ABRK3457 (1-2390)
QY 240 ValLeuValHisCysAlenulaGlyTlleSerArgSer 251
|||||
Db 1216 GTCTTGATACATTGCTTGCGCTGCGCATTTACCCGCTCA 1251

RESULT 50
ACCA6774
ID ACCA6774 standard; cDNA; 2390 BP.
XX ACCA6774;
AC
XX
XX 05-JUN-2003 (first entry)
DT
XX
XX
DE Human COPD related protein encoding cDNA SEQ ID NO:25.
XX
XX Human; chronic obstructive pulmonary disease; COPD; chronic lung disease;
XX gene; ss.
XX
XX Homo sapiens.
OS
XX WO200297127-A2.
PN
XX
XX 05-DEC-2002.
PD
XX
XX 28-MAY-2002; 2002MO-EP005835.
PF
XX
XX 31-MAY-2001; 2001GB-00013266.
PR
XX
XX (PARB ) BAYER AG.
PA
XX
XX Oellere N, Gehrmann M, Kalabis H, Hall R, Schulze T, Kroegel C;
PI WPI; 2003-140492/13.
XX
XX P-PSDB; ABP96803.
DR

```

XX Predicting, diagnosing or prognosing chronic lung disease, by detecting a
PT chronic obstructive pulmonary disease (COPD) gene in a biological sample.
XX

PS Claim 8; Page 113-114; 214pp; English.

XX
CC The present invention describes a method for predicting, diagnosing or
CC prognosing chronic lung disease by detecting a chronic obstructive
CC pulmonary disease (COPD) gene related polynucleotide (see ACC46750 to
CC ACC46777, which encode the COPD related proteins in ABP96779 to
CC ABP96806). The method is useful for predicting, diagnosing or prognosing
CC chronic lung disease in a biological sample. The COPD genes and proteins
CC encoded by them from the present invention (I) can be used for treating
CC or preventing chronic lung disease in a mammal. (I) can be used in an
CC animal model for determining the efficacy, toxicity, or side effects of
CC treatment with (I), and determining the mechanism of action of (I).
CC ACC46778 to ACC46903 represent COPD related PCR primers and probes used
CC in an example from the present invention

XX SQ Sequence 2390 BP; 565 A; 602 C; 616 G; 607 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.279	Length:	2390
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.80%	Indels:	0
DB:	7	Gaps:	0

US-10-029-345A-109 (1-665) x ACC46774 (1-2390)

QY	240	ValLeuValHisCysLeuAlaGlyIleSerArgSer	251
DB	1216	GTCTTGATACATGCTTGCTGCGCATTTAGCGCGCTCA	1251

Search completed: June 22, 2004, 07:00:42
Job time : 764 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2004, 06:47:53 ; Search time 733 Seconds
(without alignments)
4156.104 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665

Sequence: 1 MAHEMGTGIVTERLVALLE.....LKVGSQSFSGSMETIEVS 665

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3017426 seqs, 2290544650 residues

Word size: 1

Total number of hits satisfying chosen parameters: 820462

Minimum DB seq length: 1419

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.epool/US10029345/rnatc_21062004_164402_29793/app.query.fasta_1.839
-DB=Published Applications NA -QFWT=fastcap -SUFFIX=oligo.rnpb -MINMATCH=0.1
-LOOPCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human0.cgi -LIST=75 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=50 -MODE=LOCAL -OUTWMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=4419
-MAXLEN=200000000 -USER=US10029345_QCGN_1_1221_Qrnatc_21062004_164402_29793
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60
-FAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result Query *
No. Score Match Length DB ID Description

1	472	71.0	1998	9	US-09-816-494-3	Sequence 27, App1
2	472	71.0	1998	16	US-10-377-072-27	Sequence 3, App1
3	472	71.0	2102	16	US-10-094-749-673	Sequence 673, App
4	472	71.0	2732	13	US-10-168-506-2	Sequence 2, App1
5	472	71.0	2966	13	US-10-296-115-520	Sequence 520, App1
6	472	71.0	3059	17	US-10-257-026-1	Sequence 1, App1
7	472	71.0	3332	9	US-09-964-277-20	Sequence 20, App1
8	472	71.0	3496	9	US-09-964-277-1	Sequence 1, App1
9	472	71.0	3544	9	US-09-816-494-1	Sequence 1, App1
10	472	71.0	3544	16	US-10-377-072-25	Sequence 25, App1
11	472	71.0	3625	13	US-10-425-114-26234	Sequence 26234, App
12	472	71.0	3766	13	US-10-343-357-17	Sequence 17, App1
13	472	71.0	4790	17	US-10-648-593-115	Sequence 115, App1
14	408	61.4	2200	13	US-10-072-012-255	Sequence 255, App1
15	304	45.7	2071	13	US-10-072-012-257	Sequence 257, App1
16	178	26.8	1916	16	US-10-108-260A-2429	Sequence 2429, App1
17	20	3.0	2453	14	US-10-005-858-1	Sequence 1, App1
18	20	3.0	2476	13	US-10-220-120-28	Sequence 28, App1
19	12	1.8	1848	9	US-09-925-299-69	Sequence 69, App1
20	12	1.8	1848	10	US-09-925-299-69	Sequence 69, App1
21	12	1.8	1952	16	US-10-199-221-11	Sequence 11, App1
22	12	1.8	2039	13	US-10-072-012-265	Sequence 265, App1
23	12	1.8	2104	12	US-10-152-319A-2041	Sequence 2041, App1
24	12	1.8	2104	16	US-10-388-934-654	Sequence 654, App1
25	12	1.8	2109	16	US-10-305-720-1135	Sequence 270, App1
26	12	1.8	2109	16	US-10-305-720-1135	Sequence 1135, App1
27	12	1.8	2109	17	US-10-641-643-946	Sequence 946, App1
28	12	1.8	2390	13	US-10-342-887-698	Sequence 698, App1
29	12	1.8	2390	13	US-10-172-118-698	Sequence 698, App1
30	12	1.8	2390	15	US-10-175-523-183	Sequence 183, App1
31	12	1.8	2390	16	US-10-199-221-4	Sequence 4, App1
32	12	1.8	2390	17	US-10-717-597-4	Sequence 4, App1
33	12	1.8	2649	16	US-10-133-937-48	Sequence 48, App1
34	12	1.8	2649	16	US-10-159-563-48	Sequence 48, App1
35	12	1.8	2650	15	US-10-184-832-1	Sequence 1, App1
36	12	1.8	2792	10	US-09-971-392-69	Sequence 69, App1
37	12	1.8	2794	14	US-10-044-090-330	Sequence 330, App1
38	12	1.8	7001	16	US-10-199-221-12	Sequence 12, App1
39	10	1.5	165961	16	US-10-085-117-241	Sequence 241, App1
40	9	1.4	1571	15	US-10-006-288-423	Sequence 423, App1
41	9	1.4	1619	9	US-09-736-457-801	Sequence 801, App1
42	9	1.4	1619	9	US-09-902-941-801	Sequence 801, App1
43	9	1.4	1619	9	US-09-849-626-801	Sequence 801, App1
44	9	1.4	1619	13	US-10-283-017-801	Sequence 801, App1
45	9	1.4	1619	15	US-10-017-754-801	Sequence 801, App1
46	9	1.4	1619	15	US-10-113-872-801	Sequence 801, App1
47	9	1.4	1632	10	US-09-971-392-5	Sequence 5, App1
48	9	1.4	1685	14	US-10-002-600-51	Sequence 51, App1
49	9	1.4	1702	13	US-10-342-887-1034	Sequence 1034, App1
50	9	1.4	1702	13	US-10-172-118-1034	Sequence 1034, App1
51	9	1.4	1908	16	US-10-388-934-61	Sequence 61, App1
52	9	1.4	1912	16	US-10-388-934-60	Sequence 60, App1
53	9	1.4	1913	13	US-10-116-802-31	Sequence 31, App1
54	9	1.4	1998	15	US-10-103-313-648	Sequence 648, App1
55	9	1.4	2000	9	US-09-969-708-276	Sequence 276, App1
56	9	1.4	2000	9	US-09-919-497-10	Sequence 10, App1
57	9	1.4	2000	9	US-09-880-107-3774	Sequence 3774, App1
58	9	1.4	2000	10	US-09-960-706-1082	Sequence 1082, App1
59	9	1.4	2000	13	US-10-240-425-1545	Sequence 1545, App1
60	9	1.4	2000	16	US-10-305-720-1291	Sequence 1291, App1
61	9	1.4	2015	16	US-10-341-434-92	Sequence 92, App1
62	9	1.4	2064	9	US-09-736-457-825	Sequence 825, App1
63	9	1.4	2064	9	US-09-902-941-825	Sequence 825, App1
64	9	1.4	2064	9	US-09-849-626-825	Sequence 825, App1
65	9	1.4	2064	13	US-10-283-017-825	Sequence 825, App1
66	9	1.4	2064	15	US-10-017-754-825	Sequence 825, App1
67	9	1.4	2109	9	US-09-736-457-826	Sequence 826, App1
68	9	1.4	2109	9	US-09-902-941-826	Sequence 826, App1
69	9	1.4	2109	9	US-09-849-626-826	Sequence 826, App1
70	9	1.4	2109	13	US-10-283-017-826	Sequence 826, App1
71	9	1.4	2109	15	US-10-017-754-826	Sequence 826, App1
72	9	1.4	2109	15	US-10-017-754-826	Sequence 826, App1

73	9	1.4	2109	15	US-10-113-872-826	Sequence 826, App
74	9	1.4	2170	9	US-09-925-302-204	Sequence 204, App
75	9	1.4	2170	13	US-09-925-302-204	Sequence 204, App

ALIGNMENTS

RESULT 1

```
US-09-816-494-3
; Sequence 3, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIORITY FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-816-494-3
```

Alignment Scores:

Pred. No.:	0	Length:	1998
Score:	472.00	Matches:	663
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	70.98%	Indels:	4
	9	Gaps:	0

US-10-029-345A-109 (1-665) x US-09-816-494-3 (1-1998)

```
QY 1 MetAlHhGluMetIleGlyThrGlnIleValThGluArgLeuValAlaLeuEnglu 20
Db 1 ATGGCCCATGAGATGATTTGGAACTCAATTTGTTACTGAGAGTTGGTGGCTCTGCTGGAA 60
QY 21 SerGlyThGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db 61 AGTGGAAACGGAAGAAAGTGTGCTAATTGATGACCGGCATTTGTGGAATACATCATCC 120
QY 41 HisIleLeuGluAlaIleAsnIleAsnIleSerLysLeuMetLysArgLeuGlnGln 60
Db 121 CACATTTTGAAGCCATTATATCACTGCTCCAAAGCTTATGAAAGCGAAGTTGCACACG 180
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysValAspIleAsp 80
Db 181 GACAAAGTGTATATTACAGAGCTCAGCATTCAGCGAAACATTAAGTTGACATTGAT 240
QY 81 CysSerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSer 100
Db 241 TGCAGTCAGAAAGTTGTAGTTTACATCAACAGCTCCCAAGATGTGCTCTCTCTTCA 300
QY 101 AspCysPheLeuThrValLeuLeuGluLysLeuGluLysSerPheAsnSerValHis 120
Db 301 GACGTGTTTCTACCTGACTTCTGGGTAACTGGAGAAAGAGCTTAACTCTGTTACCTG 360
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGluGlyLysSer 140
Db 361 CTTCGAGGTGGGTGTGCTGAGTGTCTCTGTTGTTCCCTGGGCTCTGTGAAAGAAATCC 420
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db 421 ACTCTAGTCCCTACCTGATTTCTCAGCCTTGCTTACCTGTGCGCAACATTTGGGCAAC 480
QY 161 ArgIleLeuProAsnLeuTyrLeuGluLysGlnArgAspValLeuAsnLeuGluLeu 180
Db 481 CGAATTCTTCCCAATCTTATCTTGGCTGCGCAGAGATGTCTCCACAAAGAGCTGAT- 539
```

```
QY 181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProLysProAsp 200
Db 540 GCAGAGAAATGGGATGGTATGTGTAAATGCCAGCAA-TACCTGTCCAAAGCCTGACT 598
QY 200 helIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle 220
Db 599 TTATCCCCAGTCTATTTCCGCGTGGCCGTGTAATGACAGCCTTTGTGAGAAATTT 658
QY 220 euProTribLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCys 240
Db 659 TGCCTGTGTGACAAATCAGTAGTTCTTGAAGAAAGCAAAAGCCTCCATGATGTG 718
QY 240 alLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 260
Db 719 TTCTAGTCACATGTTTACCTGGGATCTCCGCTCGCAGCATGCTATCCCTACATCA 778
QY 260 eLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluLysArgProT 280
Db 779 TGAAGAGATGACATGCTTTAGATGAAGCTTACAGATTTGTAAAGAAAAAGACCTA 838
QY 280 hrIleSerProAsnPheAsnPheLeuGluGlnLeuLeuAspTyrGluLysLysIleLys 300
Db 839 CTATATCTCCAAACTTCAATTTTCTGGGCCAACTCCTGACTATGAGAAAGATTAAGA 898
QY 300 snGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisIleuGluLysProA 320
Db 899 ACCAGACTGAGACATCAGGGCCAAAGCAAACTCAAGCTGCTGACCTGGAGAGCCAA 958
QY 320 snGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProP 340
Db 959 ATGAACCTGTCCCTGCTGTCTCAAGAGGTGACAGAAAGGAGACGCTCTGCTCAC 1018
QY 340 roCysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSer 360
Db 1019 CCTGTGCCAAGCTGTGTAACCTCAGAGGACAGAGCAAAAGGCCCTGTCATCCGCCAGCG 1078
QY 360 alProSerValProSerValGlnProSerLeuLeuAspSerProLeuValGlnAla 380
Db 1079 TGCCAGAGGTGCCAGGCTGAGCGCTGCTGTTAAGGACAGCCCTGTGTACAGCGCG 1138
QY 380 euSerGlyLeuHisIleLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArg 400
Db 1139 TCAGTGGGCTGCACCTGTCCGCGACAGCGCTGGAAGACAGCATTAAGCTTCAAGCTTCT 1198
QY 400 heSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGly 420
Db 1199 TCTCTGTGATATCAATACAGTTTCAATATTCAGCCAGCATGGCAGCATCTTACATGGCT 1258
QY 420 heSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThra 440
Db 1259 TCTCTCATCAAGAAAGCTTTGGAAATCACTCAAACTTCCACTACTCTGGAGAGGACCA 1318
QY 440 snLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSerP 460
Db 1319 ACAAGCTATGCCAGTTCTCCCTGTTCCAGGAATATCGAGAGACATCCCGAAACAGTCT 1378
QY 460 roAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSer 480
Db 1379 CTAGTAAAGGAGAGGACAGCATCCCAAGAGCTGAGACCGCCAGAGCTTTCAGACAGCC 1438
QY 480 InSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
Db 1439 AGAGCAAGCATTTGCTTCCGTGAGAACAGCAGCGAGTGGCACCGCCAGAGGTCTCCCTT 1498
QY 500 euSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPhe 520
Db 1499 TATCTCCATGCTGACAGAGGAGCGGTGAGAGCAATTAACACACACACTTCTCTTGG 1558
QY 520 LysLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysLysTyrP 540
Db 1559 GCTTTTCCACAGCCAGCAGCATCTTCAGAAAGTGTGCTGCGCTGAGGCTTAAAGGCGTGGC 1618
```

```

QY 540LSErAsp1leuEnu1APrOGInThSerThProSerLeuThrSerSerTrpTyPha 560
Db 1619 ATCCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGCTGSTATTTTG 1678
QY 560 1aThrG1uSerSerH1sPheTySerSera1a1eTyrg1yG1Ser1a1SerTyrs 580
Db 1679 CCACAGAGrCTrCACACTTCTACTCTGCTCAGCCACTTACAGGAGCAGTCCAGTTACT 1738
QY 580 eR1lTySerCySsSerG1nLeuProThCySg1YAapG1nValTySerValArGA 600
Db 1739 CTGCCTACAGCTGCAGCCAGCTGCACCTTGCCGAGACCAAGCTATTCTGTGTCGAGGC 1798
QY 600 rGg1llyProSerAspArGAlaAspSerArGArSerTrpH1sg1uG1uSerProPhG 620
Db 1799 GGCAGAAgCAGTGCAGAGCTGAGCTGCTCCGCGGAGCTGCAGTGAAGAGAGCCCTTTG 1858
QY 620 1uLyG1nPhelysArGArSerCyG1MeG1nUheG1yG1uSer11eMetSerG1uA 640
Db 1859 AAAACAGrTTAAACCCAGAACTGCCAAATGGAATTGGAGAGACATCATGTCAAGA 1918
QY 640 sNArSerArG1uG1uEnG1yLySValG1ySerG1nSerSerPheSerG1ySerMeG 660
Db 1919 ACAGGTCAcGGAGAGAGCTGGGGAAAGTGGGAGTGCAGTCTAGCTTTTCGGGAGCATG 1978
QY 660 1u11e11eG1uValSer 665
Db 1979 AAATCATGTAGGCTCTCC 1995

RESULT 2
US-10-377-072-27
; Sequence 27, Application US/10377072
; Publication No. US20040009501A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Logen, Thomas Joseph
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Chun, Miyoung
; APPLICANT: Tsai, Pong-Ying
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
; TITLE OF INVENTION: 36692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MP103-0180NMIM
; CURRENT APPLICATION NUMBER: US/10/377,072
; CURRENT FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: US 09/895,860
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,370
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/723,806
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/843,297
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,801
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/815,419
; PRIOR FILING DATE: 2001-03-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1998

```

	TYPE: DNA	ORGANISM: Homo Sapiens
	FEATURE:	
	NAME/KEY: CDS	
	LOCATION: (1) ... (1998)	
	US-10-377-072-27	
Alignment Scores:		
Pred. No.:	0	Length: 1998
Score:	472.00	Matches: 663
Percent Similarity:	99.40%	Conservative: 0
Best Local Similarity:	99.40%	Mismatches: 2
Query Match:	70.98%	Indels: 4
DB:	16	Gaps: 0
US-10-029-345A-109 (1-665) x US-10-377-072-27 (1-1998)		
QY	1	MetAlaHISgluMetIleGlyThrGlnIleValThrGluArgPheValAlaLeuLeuGlu 20
Db	1	ATGGCCCATGAGATGATTTGGAAGCTCAAAATTTGTTACTGAGAGAGTGGTGGCTGCTGGAA 60
QY	21	SerGlyThrGlnLysValLeuLeuIleAspSerArgProPheValGluTyraThrSer 40
Db	61	AGTGAACGGAAAGAGCTGCTGCTAATTGATTGATACCGCCCACTTGTGCAATCAATACATCC 120
QY	41	HisIleLeuGlnAlaIleAsnIleAsnGlySerLysLeuMetLysArgLysGlnGln 60
Db	121	CACATTTTGAAGCCATTATATCACTGCTCCAGCCTTATGAGAGCAAGGTTGCAACAG 180
QY	61	AspLysValLeuIleThrGlnLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
Db	181	GACAAAGTGTATATACAGAGCTATCCAGCATTCAGCAAAACATTAAGTTGATCATTTGAT 240
QY	81	CysSerGlnLysValValValTyraAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db	241	TGCAGTCAGAAAGTGTGATTATTCAGATCAAGCTCCCAAGATGTTCTCTCTCTCTTCA 300
QY	101	AspCysPheLeuThrValLeuLeuGlyLysLeuGlnLysSerPheAsnSerValHisLeu 120
Db	301	GACTGTTTCTCAGCTGACTTCTTGGGTAACTGGAGAAAGCTTCAACTCTGTTCACTG 360
QY	121	LeuAlaGlyGlyPheAlaGlnPheSerArgCysPheProGlyLeuGlyGlnGlyLysSer 140
Db	361	CTTGCAGAGGGGTTTGTGAGATTCTCTCGTTGTTTCCCTGAGCCTCTGTGAAGAAATCC 420
QY	141	ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db	421	ACTCTAGTCCCACTGCACTTCTCAGCCCTTCACTGTTGCCAACATTTGGGCCCAAC 480
QY	161	ArgIleLeuProAsnLeuTyrlLeuGlyCysGlnArgAspValLeuAsnLysGlnLeuIle 180
Db	481	CGAATTTCTCCAACTTTATCTTGGCTGCCAGCGAGATGTCTCAACAAAGAGCTGAT- 538
QY	181	GlnGlnAsnGlyIleGlyTyraValLeuAsnAlaSerTyraThrCysProLysProAsp 200
Db	540	GCAGCAGAGATGGGATTGGTTATGTGTTAATGCAGCAAA-TACCTGTCCAAAAGCTGACT 598
QY	200	HeIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlnLysIleLeu 220
Db	599	TTATCCCCGAGCTCATATTTCTCGCGGTGCTCGTGAATATACAGCTTTTGTGAAGAAATTT 658
QY	220	euProTyrLeuAspLysSerValAspPheIleGlnLysValLysValAsnSerAsnGlyCys 240
Db	659	TGCCCTGTTGGCAAAATAGTAAATTCATTGAGAAACCAAAAGCTCCAAATGAGATGTG 718
QY	240	AlaLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlIle 260
Db	719	TTCTAGTGCACAGTTTAACTGGGATCTCCCGCTCCGCCACATCGCTATCGCTCAATCA 778
QY	260	eLysArgMetAspMetSerLeuAspGlnAlaTyraPheValLysGlnLysArgPro 280
Db	779	TGAAGAGATGACATGCTTTAGATGAAGCTTACAGATTTTGGAAAGAAAGAAAGACTTA 838

QY 280 hr11eSerProasnPhenPheLeuGlyGlnLeuLeuaspTyrGluLysLeu11eLysA 300
DB 839 CTATATCTCCAAACTTCAATTTTCTGGGCCCACTCCCTGAGCTATGAGAAAGATTAGAA 898
QY 300 snGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLysPheA 320
DB 899 ACCGAGCTGAGCATCAGGGCCAAAGAACCAACTCAAGCTGCTGCACTGGAGAGCCAA 958
QY 320 snGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProp 340
DB 959 ATGAACTGTCTCCTGCTGTCTCAGAGGGGTGAGCAAGAAAGCGAGACGCTCCAGTCCAC 1018
QY 340 rOCysAlaaspSerAlaThrSerGluAlaIaIaGlyInaGProValH1eProAlaSerY 360
DB 1019 CTTGTGCCGACTGTGCTACTCAGAGGCAAGAGCAAAAGCCCGTGCATCCCGCCAGCG 1078
QY 360 a1ProSerValProSerValGlnProSerLeuLeuGluaspSerProLeuValGlnAlaI 380
DB 1079 TGCCCAAGCGGCCAGGCTGACAGCCGCTGTAGAGGACAGCCGCTGTG7ACAGGCGC 1138
QY 380 euSerGlyLeuH1sLeuSerAlaaspArgLeuGluaspSerLeuLysLeuLysArgSerP 400
DB 1139 TCAGTGGGCTGCACCTGTCCCGCAGACAGGCTGGAGAGACAAATAAAGCTCAAGCTTCCT 1198
QY 400 heSerLeuasp11eLysSerValSerTyrSerAlaSerMetAlaIaIaSerLeuH1sGlyP 420
DB 1199 TCTCTCTGAGATCAATCAGTTTCAATTCAGCCAGCATGGCAGCATCTTATCAGGCT 1258
QY 420 heSerSerSerGluaspAlaLeuGluThrTyrTyrLysProSerThrThrLeuaspGlyThA 440
DB 1259 TCTCCATCAGAAAGATGCTTGGAAATCAAAACCTTCCATCTCAGATGGGAGCA 1318
QY 440 snLysLeuCyGlnPheSerProValGlnLysLeuSerGluGlnThrProGluThrSerP 460
DB 1319 ACAAGTATGCCAGCTTCTCCCTGTTCAAGAACTATCCGAGAGCACTCCCGAAACCAAGTC 1378
QY 460 rOAspLysGluGluAlaSer11eProLysLysLeuGlnThrAlaArgProSerAspSerG 480
DB 1379 CTGATTAAGAGAGGAGGACAGCATCCCAAGAGCTGAGACCGCCAGGCTTACAGACGCC 1438
QY 480 1nSerLysArgLeuH1sSerValArgThrSerSerSerGlyThrAlaGlnaArgSerLeu 500
DB 1439 AAGAGCAAGGATTCATTCGGGTGAGAACCAAGCAGCTGACCGCCAGAGAGTCCCTTT 1498
QY 500 euSerProLeuH1sArgSerGlySerValGluaspAsnTyrH1sThrSerPheLeuPheG 520
DB 1499 TATCTCCATCGATCGAAGTGGGAGCGTGGAGGACATTAACAACACAGCTTCTTTTCG 1558
QY 520 1yLeuSerThrSerGlnGlnH1sLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrP 540
DB 1559 GCGTTTCCACAGCCGACGACGACCTTCACGAGCTGTGCTGGGCTTAAAGGCGCTGGC 1618
QY 540 1eSerAsp11eLeuAlaProGlnThrSerThrProSerLeuH1sSerSerTyrPyrPheA 560
DB 1619 ACTCGGATATCTTGGCCCCCAGACCTTCAACCTTCCCTGACCAAGCTGTGATTTTTCG 1678
QY 560 1aThrGlnSerSerH1sPheTyrSerAlaSerAla11eTyrGlyGlySerAlaSerTyrS 580
DB 1679 CCACAGAGCTCTCACACTTCTACTCTGCTCAGCCATTCAGGAGGCGAGTGCATTTACT 1738
QY 580 eR1aTyrSerCySerGlnLeuProThrCyGlyLysGlnValLysSerValArgArgA 600
DB 1739 CTGCTTACAGCTGACGAGCGCTGCCACTTGGGAGAACCAAGTCAATTCGTGCGCCAGGC 1798
QY 600 rGg1nLysPProSerAspArgAlaaspSerArgArgSerTyrH1sGlnGlnLysSerPropheG 620
DB 1799 GGCACAAGCGCAAGTACAGAGCTGACTCGCGCGAGCTGGCATTAAGAGAGCCCTTTTCG 1858
QY 620 1uLysGlnPheLysArgArgSerCyGlnMetGluPheGlyGlnSer11eMetSerGlnA 640
DB 1859 AAAACAGATTAAACGAGAGAGCTCCCAATGAAATTTGAGAGAGCATTCATGTACAGGA 1918
QY 640 snArgSerArgGlnLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetG 660

DB 1919 AAGGCTACGGAGAGAGCTGGGGAAAGTGGGAGTCAAGTCTTTCGGGCGACATGG 1978
QY 660 1u11e11eGluValSer 665
DB 1979 AAATCATTTGAGGTCTCC 1995
RESULT 3
US-10-094-749-673
; Sequence 673, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENDU
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 08435/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 673
; LENGTH: 2102
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-673
Alignment Scores:
Pred. No.: 0
Score: 472.00
Percent Similarity: 99.10%
Best Local Similarity: 99.10%
Query Match: 70.98%
DB: 16
Gaps: 0
US-10-029-345A-109 (1-665) x US-10-094-749-673 (1-2102)
QY 1 MetAlaH1sGluMet11eGlyThrGln11eValThrGluArgLeuValAlaLeuLeuGlu 20
DB 56 ATGGCCCATGAGATGATGGAACTCAATTTGTTACTGAGAGGTTGGCTGCTGGGAA 115
QY 21 SerGlyThrGluLysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 116 AGTGAACGGAAAAAGTGGCTGATTAATGATAGCCGCGCATTTTGGAAATCAATACATCC 175
QY 41 H1s11eLeuGluAla11eAsn11eAsnCySerLysLeuLeuLysArgArgLeuGln 60
DB 176 CACATTTTGGAAACCATTAATATCAACTGCTCCAAAGCTTAATGAAGCAAGGTTGCAACAG 235
QY 61 AspLysValLeu11eThrGluLeu11eGlnH1sSerAla1yH1sLysValaAsp11eAsp 80
DB 236 GACAAAGTGTAAATTTACAGAGCTCATCCAGCATTCAGCGAAACATAAAGTTGACATTGAT 295
QY 81 CySerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100


```

Db      296 TCGAGTCAGAGAGTTGTAGTTACGATCAAGCTCCAGAGATGTTCCCTCTCTCTTCA 355
Qy      101 AAPPQVPhleuThrValLeuLeuGlyLyLeuGlnLysSerPheAnsSerValHisLeu 120
Db      356 GACTGTTTCTCACTGACTTCTGGGTTAAACTGGAGAAAGAGCTTCAACTGTTTCACTG 415
Qy      121 LeuAlaGlyGlyPheAlaGlyPheSerArgCysPheProGlyLeuGlyGlyGlySer 140
Db      416 CTTCGAGGAGGAGTTGCTGAGTCTCTGCTGTTTCCCTGGCCCTGCTGTAAGAGAAATCC 475
Qy      141 ThrLeuValProThrCysHisLeuSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db      476 ACTCTAGTCCCTACCTGCACTTCTCAGCCCTGCTTACCTGTTCCCAACATTGGGCCAAC 535
Qy      161 Arg-1LeuProAsnLeuTyrluGlyCysGlnArgAspValLeuAsnLysGlyLeu1 180
Db      536 CT-AATTTCTCCCAATCTTATCTTGCTGGCCAGAGATGTCCTCAACAGAGAGCTGAT 594
Qy      180 e-GlnGlnAsnGlyIleGlyTyrlValLeuAsnAlaSerTyrl-ThrCysProLysProAsp 199
Db      595 -GCAGAGAGATGGATGGATTGATTGTTAAATGCCAGCAA-TACCTGTCCAAAGCCTGAC 652
Qy      200 PheIleProGlySerHisPheLeuArgValProValAsnAspSerPheCysGlyLysIle 219
Db      653 TTTATCCCGAGTCTCATTTCCCTGCGCTGCTGTAAATGACAGCTTTTGTGAGAAATTT 712
Qy      220 LeuProTrpLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAnglyCys 239
Db      713 TTGCCGTGGTTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAACCTCCATGTGATGT 772
Qy      240 ValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrlle 259
Db      773 GTTCTAGTCACAGTTTACTGGGATCTCCCGCTGCCACCACTCCGTATCGCTACATC 832
Qy      260 MetLysArgMetAspMetSerLeuAspGlnAlaTyrlArgPheValLysGlyLysArgPro 279
Db      833 ATGAGAGAGATGACATGCTCTTATGATGAAGCTTACAGATTGTGAAAGAAAACAACT 892
Qy      280 ThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrlGlyLysIleLys 299
Db      893 ACTATATCTCCAAACTTCAATTTTCTGGGCCAACTCTGGAATAAGAGAAAGATTAA 952
Qy      300 AsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlyLysPro 319
Db      953 AACCAAGCTGAGACATCAGGGCCAAAGACCAAACTCAAGCTGTGACCTGGAGAGCA 1012
Qy      320 AsnGluProValProAlaValSerGlyGlyGlnLysSerGlnThrProLeuSerPro 339
Db      1013 AATGAACCTGTCCCTCTGCTCTCAGAGGTGACAGAAAGCGAGAGCCCTCAGATCCA 1072
Qy      340 ProCysAlaAspSerAlaThrSerGlnAlaAlaGlyGlnArgProValHisProAlaSer 359
Db      1073 CCTGTGCGGACCTCTGCTACCTCAGAGGAGCAGCAAAAGCCGTGATCCCGCCAG 1132
Qy      360 ValProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAla 379
Db      1133 GTGCCAGGCGTGGCCAGCGTGCAGCGCTGCTTGAAGACAGCCCGCTGTACAGGG 1192
Qy      380 LeuSerGlyLeuHisLeuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSer 399
Db      1193 CTCAGTGGGCTGACCTGTCCGACAGAGCTGGAAGACAGCAATAGCTCAAGCCGTCC 1252
Qy      400 PheSerLeuAspIleLysSerValSerTyrlSerAlaSerMetAlaIleSerLeuHisGly 419
Db      1253 TTCTCTCTGGATATCAATCAGTTTCAATTTCAGCCAGCATGGCAGCATCTTACATG 1312
Qy      420 PheSerSerSerGluAspAlaLeuGlyTyrlTyrlArgProSerThrThrLeuAspGlyThr 439
Db      1313 TTCTCTCATCAAAATGCTTTGGAAATCTACAAACCTTCCACTCTGGATGGAGAC 1372
Qy      440 AsnLysLeuCysGlnPheSerProValGlnGluLeuSerGlnThrProGluThrSer 459

```

```

Db      1373 AACAGCTATGCCAGTCTCCCTGTTGAGAACTATCGAGACAGACTCCCAACCACT 1432
Qy      460 ProAspLysGlnGlnAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSer 479
Db      1433 CCTGATTAAGAGAAAGCCAGCATCCCAAGAGCTGACATGCGAGGCTTTCAGACAC 1492
Qy      480 GlnSerLysArgLeuHisSerValArgThrSerSerSerGlyTyrlAlaGlnArgSerLeu 499
Db      1493 CAGAGCAAGCATTTGATTCGATCGATCAGAACACAGACAGATGGACCCCGAGAGTCCCT 1552
Qy      500 LeuSerProLeuHisArgSerGlySerValGluAspAsnTyrlHisThrSerPheLeuPhe 519
Db      1553 TTATCTCCACTCATCGAAGTGGAAGCGTGGAGGACAAATTATCACACCAAGCTTCTTTCC 1612
Qy      520 GlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrl 539
Db      1613 GGCCTTTCCACAGCAGCAGCAGCAGCTCAGAAAGTCTGCGCTGGCTTAAAGGCTGG 1672
Qy      540 HisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTyrlPhe 559
Db      1673 CACTGGATATCTTGACCCCGACAGCTTACCTTCCCTGACAGCAGCTGTGATTTT 1732
Qy      560 AlaThrGluSerSerHisPheTyrlSerAlaSerAlaIleTyrlGlySerAlaSerTyrl 579
Db      1733 GCCACAGAGTCTCTACACTTACTGCTGCTCAGGCATCTACGAGAGCAGTGCACATTAC 1792
Qy      580 SerAlaTyrlSerCysSerGlnLeuProThrCysGlyAspGlnValTyrlSerValArgArg 599
Db      1793 TCTGCTCAGCTGAGCAGCCAGCTGCCCACCTTCCGAGACCAAGTCAATCTGTGCCAG 1852
Qy      600 ArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGlnGluSerProPhe 619
Db      1853 CGGCAAGACCAAGTACAGACTGACTGCGGAGAGCTGGATGGAAGAGAGAGCCCTTT 1912
Qy      620 GlyLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGlu 639
Db      1913 GAAAAGCATTTAAACGCAAGAGCTGCCAAATGGAAATTTGGAGAGCATCATGTCAAG 1972
Qy      640 AsnArgSerArgGlnGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMet 659
Db      1973 AACAGGTCACGGGAAGAGCTGGGGAAGTGGGACAGTCACTTTCGGGACAGCATG 2032
Qy      660 GlnIleIleGlnValSer 665
Db      2033 GAATCATTAGAGTCTCC 2050

RESULT 4
US-10-168-506-2
; Sequence 2, Application US/10168506
; Publication No. US20040053229A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: HILL, RON
; APPLICANT: FLANAGAN, PETER
; TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
; FILE REFERENCE: 038602/1351
; CURRENT APPLICATION NUMBER: US/10/168,506
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: PCT/US00/34736
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-168-506-2
Alignment Scores:

```

Pred. No.: 0
Score: 472.00
Percent Similarity: 99.40%
Best Local Similarity: 99.40%
Query Match: 70.98%
DB: 13
Length: 2732
Matches: 663
Conservative: 0
Mismatch: 2
Indels: 4
Gaps: 0

US-10-029-345a-109 (1-665) x US-10-168-506-2 (1-2732)

QY 1 MetAlHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 538 ATGGCCCATGAGATGATGGACTCAAAATGTTACTGAGAGGTGGTGGCTGTGCTGAA 597
QY 21 SerGlyThrGluHisValLeuLeuIleAspSerArgProPheValGluIleThrAsnThrSer 40
DB 598 AGTGAACCGAAGAAAGTGGCTGTAATGATGAGCCGGCCATTTGTGAATAACATATCATCC 657
QY 41 HisIleLeuGluValIleAsnIleAsnCySerIleLeuMetLysArgArgLeuGlnGln 60
DB 658 CACATTTTGGAGCCATTAATATCACTGCTCCAACTTATGAGGAAAGTTGGCAACAG 717
QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80
DB 718 GACAAAGTGTATATACAGAGCTCATCCAGCAATTCAGCGAAACATAGGTTGACATTGAT 777
QY 81 CySerGluLysValValIleValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 778 TGCAGTCAGAGAGCTTTAGTTTACGATCCAAAGCTCCCAAGATGTTGCCCTCTCTCTTCA 837
QY 101 AspCySerPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
DB 838 GACTGTTTCTCACTGTACTTCTGGGTAACTGGAGAAGAGCTCAACTCTGTTACCTGG 897
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCySerPheProGlyLeuCySerGluGlyLysSer 140
DB 898 CTTGACAGGAGGTGGTGGCTGAGTTCCTCGTTGTTCCCTGGGCTCTGTGAAGAAATCC 957
QY 141 ThrLeuValProThrCySerIleSerGlnProCySerLeuProValAlaAsnIleGlyProThr 160
DB 958 ACTCTAGTCCCTACCTGCACTTCTGACCTTGTCTTACCTGTTGCCAACATGGGCCAAC 1017
QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCySerGlnArgAspValLeuAsnLysGluLeuIle 180
DB 1018 CCAATTTCTTCCCAATCTTATCTTGGCTGCCAGCGAGATGCTCTCAACAAGAGCTGAT - 1076
QY 181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCySerProLysProAsp 200
DB 1077 GCAGCAGAAATGGATGTTATGTGTTAATGCCAGCA-TACCTGTCCAAAGCCTGACT 1135
QY 200 HeIleProGluSerHisPheLeuArgValProValAsnAspSerPheCySerGluLysIle 220
DB 1136 TTATCCCGCAGTCTCATTTCTCGCGGTGCCCTGTGAATACAGCTTTTGTGGAAATTT 1195
QY 220 euProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCySer 240
DB 1196 TGCCCTGGTGGCAAAATCAGTAGATTCATTGAGAAAGCAAAAGCCTCCATGATGAG 1255
QY 240 AlLeuValHisCySerLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 260
DB 1256 TTCAGTGTGACTGTTAGCTGGAGATCTCCGCTCCGCCCAACATCCCTTCGCTTCATCA 1315
QY 260 eLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProT 280
DB 1316 TGAAGAGATGACATGCTCTTTTATGATGAAGCTTACAGATTTGTGAAGAAAGAAAGACCTA 1375
QY 280 hTlleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLysIleLys 300
DB 1376 CTATATCTCCAAACTTCAATTTTCTGGGCCCAACTCCTGGACTATAGAAAGAGATTAGA 1435
QY 300 snGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisIleGluLysPro 320
DB 1436 ACCAGACTGAGCATCAAGGCCAAAGAGCAAACTCAAGCTGCTGACCTGGAGAGGCCAA 1495

QY 320 snGlnProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerPro 340
DB 1496 ATGAACCTGTCCCTCTCTGCTCAGAGGGTGGACAGAAAGCGACAGCGCCCTCACTCAC 1555
QY 340 roCyAlaAspSerAlaThrSerGluAlaAlaGlyLysArgProValHisProAlaSer 360
DB 1556 CCTGTGCCGACTCTCTACCTCAGAGGACGACGAGCAAAAGGCCCTGTGATCCCGCAGCG 1615
QY 360 AlProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAla 380
DB 1616 TGCCAGAGGTGCCAGCTGACCGCTGAGCCGTGCTGTTAGAGACAGCCCGCTGTAAGCGGC 1675
QY 380 euSerGlyLeuHisIleSerSerAlaAspArgLeuGluAspSerAsnLysLysArgSer 400
DB 1676 TCAGTGGCTGACCTGTCCCGACAGCTGGAGAGACAGCAATTAAGCTCAAGCGTTCT 1735
QY 400 heSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyP 420
DB 1736 TCTCTGTGATATCAATCAAGTTTCAATTCAGCCAGCATGGCAGCATCTTACATGGCT 1795
QY 420 heSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThra 440
DB 1796 TCTCTCATCAGAAAGATGCTTGGAACTACAAACCTTCCACTACTCTGGATGGAGCA 1855
QY 440 snLysLeuCySerGlnPheSerProValGlnGluLeuSerGluIleThrProGluThrSerP 460
DB 1856 ACAGCTATGCGAGTTCCTCCCTGTCAGAACTATCGAGACAGCTCCCAAAACCAATC 1915
QY 460 roAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSer 480
DB 1916 CTGATTAAGAGAGAGCCAGCATCCCAAGAAAGCTGACAGCTCCAGAGCTTCAAGACGCC 1975
QY 480 InSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
DB 1976 AGAGCAAGCATGTGATTCGCTCAGAACCAAGCAGAGGACCGCCCAAGAGTCCCTTT 2035
QY 500 euSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPhe 520
DB 2036 TATCTCCACTCATCGAATGAGAGCGTGGAGACAAATTACCAACAGCTTCTTTTCG 2095
QY 520 ILeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrP 540
DB 2096 GCCTTTCACAGCAGCAGACGACCTCAAGAACTGCTGGCTGGCCTTAAGGGCTGGC 2155
QY 540 iSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPhe 560
DB 2156 ACTGGATATCTGGCCCCCAGACCTTACCCCTTCCTGACAGAGCTGTATTTTG 2215
QY 560 laThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyr 580
DB 2216 CCAGAGTCTTCACTTCTACTCTGCTCAGCACTTACGAGGCGAGTCCAGTTTACT 2275
QY 580 eAlaTyrSerCySerGlnLeuProThrCySerGlyAspGlnValTyrSerValArgArg 600
DB 2276 CTGCTACAGCTGACAGCAGCTGCCACTTGCAGAGCAAAAGTCTATTCGTGGCAGGC 2335
QY 600 rGlnLysProSerAspArgAlaAspSerArgAspSerTrpHisGluGluSerProPhe 620
DB 2336 GGCAGAAAGCCAAAGAGACAGACTACCTGCGCGAGACTGGCAGAAAGAGGCCCTTTG 2395
QY 620 LuLysGlnPheLysArgArgSerCySerGlnMetGluPheGlyGluSerIleMetSerGlu 640
DB 2396 AAAAGCAGTTTAAACGAGAAAGCTCCAAATGAAATTTGAGAGAGCATCTGTCAGAGA 2455
QY 640 snArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetG 660
DB 2456 ACAGCTACGGAGAGAGCTGGGAGAAAGTGGCAGTCACTGACTTTCCGGCAGCATGG 2515
QY 660 luIleIleGluValSer 665
DB 2516 AATCATTTGAGTCTCC 2532

RESULT 5

QY 480 InsertYsArGLeuHIsSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuI 500
DB 1565 AGAGCAAGCATTCGATTCGGTGTGAGAACGACGACGATGGACACCGCCAGAGGTCCCTTT 1624
QY 500 euSerProLeuHIsArGSerGlySerValGluAapAntYrHIsrThrSerPheLeuPheG 520
DB 1625 TATCTCCACTGCATCGAAGTGGAGCGTGGAGCAATTCACACACAGCTTCCTTTGG 1684
QY 520 lyeuSerThrSerGlnGlnHIsLeuThrIysSerAlaGlyLeuGlyLeuIyTrpH 540
DB 1685 GCGTTTCCACACACGACGACACACTCAGCAAGTCTGCTGGCGCTTAAAGCGCTGGC 1744
QY 540 lIsSerAplIleuAlaProGlnThrSerThrProSerLeuThrSerSerTrpYrPheA 560
DB 1745 ACTCGGATATCTTGCGCCCGCCCAACCTTACCCCTTCCTGACGACGACGCTGATTTTG 1804
QY 560 lATHrGluSerSerHIsPheTySerAlaSerAlaIleTyGlyGlySerAlaSerTyS 580
DB 1805 CCAAGAGTCTCTACACTTCTACTGCTGCTCAGCCATCTACGAGGAGGAGTCCAGTTACT 1864
QY 580 exrAlaTySerCySerSerGlnLeuProThrCyGlyAapGlnValTySerValArgArg 600
DB 1865 CTGCTTACAGCTCAGCGACGCTGCCACTTGCAGAGCCAAAGTCTATTCTGTGCGCAGGC 1924
QY 600 rGgInIysProSerAaPArgAlaAapSerArgArGSerTrpHIsGluGluSerProPheG 620
DB 1925 GGGAGAAAGCCAAAGTGAAGAGTGAAGTCTGCGGCGAGCTGGCAATGAAGAGGCCCTTGG 1984
QY 620 lulysGlnPheTyArGArgSerCyGlnImetGluPheGlyGluSerIleMetSerGluA 640
DB 1985 AAAAGCAGTTTAAACGAGAGAGCTGCCAATGGAATTTGGAGAGAGATCATGTCAAGA 2044
QY 640 smArGSerArgGluGluLeuGlyIyValGlySerGlnSerPheSerGlySerMetG 660
DB 2045 ACAGGTCAACGGGAAGCTGGGGAAGTGGGCAATGCTGACTTTTCGGGCAAGATGG 2104
QY 660 lulIleGluValSer 665
DB 2105 AAATCATTTAGAGTCTCC 2121

RESULT 7
US-09-964-277-20 : Sequence 20, Application US/09964277
: Patent No. US20020137170A1
: GENERAL INFORMATION:
: APPLICANT: Luche, Ralf M.
: ACT: Wei, Bo
: TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
: FILE REFERENCE: 200125.434
: CURRENT APPLICATION NUMBER: US/09/964,277
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 20
: LENGTH: 3332
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-964-277-20

Alignment Scores:
Pred. No.: 0 Length: 3332
Score: 472.00 Matches: 472
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.98% Indels: 0
DB: 9 Gaps: 0

US-10-029-345a-109 (1-665) x US-09-964-277-20 (1-3332)

QY 194 ThrCyProIysProAapPheIleProGluSerHIsPheLeuArgValProValAaP 213
DB 977 ACCTGTCAAAAGCTGACTTATCCCGAGTCTCATTTCTGTGTGCTGTGAATGAC 1036

QY 214 SerPheCyGluIyValIleuProTrPleuAapIySerValAapPheIleGluIyAla 233
DB 1037 AGCTTTTGAGAAATTTTGGCGTGTGGACAAATCGATGATTTTCATGGAAAGCA 1096
QY 234 lyeAlaSerAaGlyCyValIleuValHIsCyValLeuAlaGlyIleSerArgSerAlaTrH 253
DB 1097 AAAGCTCCAAATGAGATGTGTTCAGTGCAGCTGTTAGTGGATTCGCCGCTCCGACAC 1156
QY 254 lIleAlaIleAlaTyrlIleMetIyAaGMeAaPMeSerIleuAaGluAlaTyraGPh 273
DB 1157 ATCGCATGCTCTACATCAATGAAGAGATGACATGTCTTATGATGAAGCTTACAAATTT 1216
QY 274 ValIySerGluIyAaPProThrIleSerProAaPheAaPheLeuGlyGlnIleuAaP 293
DB 1217 GTGAAGAAAGAAAGACTCTATATATCCAACTTAATTTTCTGGCCAACTCTCGGAC 1276
QY 294 TyrGluIyLeuIleIyAaGlnIThrGlyAlaSerGlyProIySerIySerIyLeuIyLeu 313
DB 1277 TATGAGAAAGATTTAAGAACGACTGAGCATTCAGGCGCAAGAGCAAACTCAAGCTG 1336
QY 314 lIeHIsLeuGluIySerProAaGlnIProValProAlaValSerGluGlyGlnIySer 333
DB 1337 CTGCACTGAGAGAGCCAAATGACCTGTCCGTCTCAGAGGGGTGACAGAAAGC 1396
QY 334 GluThrProLeuSerProProCyAlaAapSerAlaThrSerGluAlaIleGlyIyArg 353
DB 1397 GAGACGCCCTCACTCAGCTGCTGTGCGCACTGCTCACTCAGAGGACAGCAAGCAAG 1456
QY 354 ProValHIsProAlaSerValProSerValProSerValGlnProSerLeuGluAaP 373
DB 1457 CCGGTGATCCCGCCAGCGCTGCCAGCGTCCAGCGTCCAGCGTCCGCTTTAGAGAC 1516
QY 374 SerProLeuValGlnAlaLeuSerGlyIyLeuHIsLeuSerAlaAaPArgLeuIyAaP 393
DB 1517 AGCCCGCTGTACAGCGCTGAGTGGCTGCACTGTCCGCAAGCAGGTGGAAGCAGC 1576
QY 394 AaMlyIyLeuIyAaPArgSerPheSerLeuAaPrlIyIySerValSerTySerIleAaMet 413
DB 1577 AAATGCTCAAGGATCTTCTCTGGAATCAATCAAGTTTCATATTCAGCACATG 1636
QY 414 AlaAlaSerLeuHIsGlyPheSerSerSerGluAaPAlaLeuGlnIyTyrlIyIyProSer 433
DB 1637 GCAGATCTCTTACATGCTTCTCTCATGAAAGATCTTTGGAATCATCAAACTTCC 1696
QY 434 ThrThrLeuAaPrlIyThrAaIyIyLeuCyGlnPheSerProValGlnIyLeuSerGlu 453
DB 1697 ACTACTCTGATGGAGCAAAAGCTATGCTCCAGTTCTCCCTGTTCAGGAACATATCGAG 1756
QY 454 GlnThrProGluThrSerProAaPrlIyGluAlaSerIleProIyIyIyLeuGlnThr 473
DB 1757 CAGACTCCGAAACCGCTCTATAGAGAGAGAGAGATGCCCAAGAGCTGCAAGACC 1816
QY 474 AlaArgProSerAaPArgSerIyIyAaGLeuHIsSerValArgThrSerSerGly 493
DB 1817 GCCAGGCTTTCAGACAGCCAGAGCAAGCATTCGCTCAAGAACGACGACGATGGC 1876
QY 494 ThrAlaGlnArgSerLeuLeuSerProLeuHIsArGSerGlySerValGluAaPAntYr 513
DB 1877 ACCGCCAAGGTCCTTTATCTCAGCATCGAAGTGGAGCGTGGAGGACAAATTAC 1936
QY 514 HIsThrSerPheLeuPheGlyIyLeuSerThrSerGlnGlnHIsLeuThrIySerAlaGly 533
DB 1937 CACACAGATTCCTTTTGGCTTTCACACGACGACACACCTTCGAAAGTGTGCTGGC 1996
QY 534 lIeGlyIyLeuIyGlyTrpHIsSerAaPrlIleuAlaProGlnThrSerThrProSerLeu 553
DB 1997 CTGGGCTTAAAGGCTGAGCATCGATATCTTGCCCCCAAGCTTACCCCTTCCCTG 2056
QY 554 ThrSerSerTrpTyrPheAlaThrGluSerSerHIsPheTySerAlaSerAlaIleTy 573
DB 2057 ACCAGAGCTGTGATTTTCCACAGAGTCTCACACTTCTACTCTGCTCAGGACATGAC 2116

QY 574 GlyIYSerAlaSerTyrSerAlaTyrSerCysSerGluLeuProThrCysGlyAspGln 593
DB 2117 GGAGGCAAGTGCAGTACTGCTGACAGCTGCAGCCAGCTGCCACTGGGAGACCAA 2176
QY 594 ValTyrSerValArgArgArgGlnIysProSerAspArgAlaAspSerArgArgSerTyr 613
DB 2177 GTCTATTTCTGCGCAGCGCGCAGAAAGCCAGAGTGCAGAGTCACTGCCGCGAGCTGG 2236
QY 614 HisGluGluSerProPheGluIysGlnPheIysArgArgSerCysGlnMetGluPheGly 633
DB 2237 CATGAGAGAGCCCTTTGAAAACAGCTTTAAAGCAGAAAGCTGCCAAATGAGATTGGA 2296
QY 634 GluSerIleMetSerGluAsnArgSerArgGluGluLeuGlyIysValGlySerGlnSer 653
DB 2297 GAGAGCATCATGTCAAGAGAACAGGTCAACGGAAGCTGGGAGAAAGTGGCAGTCACT 2356
QY 654 SerPheSerGlySerMetGluIleIleGluValSer 665
DB 2357 AGCTTTTCGGGACATGAAATCATTCAGGCTCC 2392

RESULT 8

US-09-964-277-1
; Sequence 1, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125 434
; CURRENT APPLICATION NUMBER: US/09/964, 277
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-964-277-1

Alignment Scores:

Pred. No.: 0 Length: 3496
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 9 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-964-277-1 (1-3496)

QY 1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB 562 ATGGCCCATGAGATGATGTAAGTCAATGTTTACTGAGAGTGTGGCTCTGCTGAA 621
QY 21 SerGlyThrGluIysValIleuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
DB 622 AGTGAAGAGGAAAAAGTCTGCTCAATGATGATACCGGCAATTTGTGAAATACAAATCATCC 681
QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerIysLeuMetIysArgArgLeuGlnGln 60
DB 682 CACATTTTGGAGCATTATATCAACGCTCCAAAGCTTATGAAAGCGAAGTTCACAG 741
QY 61 AspIysValIleuIleThrGluLeuIleGlnHisSerAlaIysHisIysValAspIleAsp 80
DB 742 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGGAAACATTAAGTTGACATTGAT 801
QY 81 CysSerGlnIysValIleValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 802 TGCACATCAGAAAGTGTGTTTACATTAAGAGCTCCCAAGATGTTGCTCTCTCTTCA 861
QY 101 AspCysPheLeuThrValIleuLeuGlnIysLeuGlnIysSerPheAsnSerValHisLeu 120
DB 862 GACTGTTTTCACACTGACTTCTGGGTAAACTGGAGAAAGACTTCACTCTGTTCACTG 921

QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyIleuCysGluIysSer 140
DB 922 CTTCGAGAGTGGGTTTGCTGAGTCTCTGCTGTTTCCCTGGGCTCTTGGAAGAAATCC 981
QY 141 ThrIleValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 982 ACTCTAGTCCCTACCTGCAATTCAGGCTTGCTTACCTGTTTGCCAAATGGGCAACC 1041
QY 161 ArgIleLeuProAsnLeuTyrIleuGlyCysGlnArgAspValIleAsnIysGluLeuIle 180
DB 1042 CGAATTTCTCCAACTTATATCTTGGCTGCGCAGCAGATGTTCTCAACAGAGCTGAT- 1100
QY 181 -GlnGlnAsnGlyIleGlyTyrValIleuAsnAlaSerTyr-ThrCysProIysProAsp 200
DB 1101 GCAGCAGATGGATGGTATGTTGTTAAATGCCAGCAA-TACGCTCCAAAGCTGACT 1159
QY 200 IleIleProGluSerIlePheLeuArgValProValAsnAspSerPheCysGluIysIle 220
DB 1160 TTATCCCGAGCTCATTTCTGCGTGCCTGTGAATGACAGCTTTGTGAGAAATTT 1219
QY 220 euProIleuAspIysSerValAspPheIleGluIysAlaIysAlaSerAsnIysCysV 240
DB 1220 TGCCTGTTGGCAAAATCAGTATTTCAATTGAGAAAGCAAAAGCTCCATGAGATGTG 1279
QY 240 AlIleValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleM 260
DB 1280 TTCTAGTGCACCTGTTTAGCTGGATCTCCGCTCCGACCATGCTATGCTTACATCA 1339
QY 260 eLysArgMetAspMetSerIleuAspGluIleTyrArgPheValIysGluIysArgPro 280
DB 1340 TGAAGAGATGAGATGCTTTTAAATGAGCTTACAAATTTGTGAAGAAAAAGACTTA 1399
QY 280 hrIleSerProAsnPheAsnPheLeuGlyIleuLeuAspTyrGluIysIleIysA 300
DB 1400 CTATATCTCCAACTTCAATTTCTGCGGCAACTCTGAGCATGTGAGAAAGATTAGA 1459
QY 300 snGlnThrGlyAlaSerGlyProIysSerIysLeuIysLeuHisIleuGluIysPro 320
DB 1460 ACCAGACTGGAGCATCGGCGCAAGAGCAAACTCAAGCTGACCTCGAAGAAAGCCTCA 1519
QY 320 snGluProValProAlaValSerGluIysGlyIysGlnIysSerGluThrProIysSer 340
DB 1520 ATGAACCTGTCCTGCTGCTCAGAGGGGTGAGCAAGAAAGCGAGCCCTTAGTCCAC 1579
QY 340 roCysAlaAspSerAlaThrSerGluAlaIleGlyIlnArgProValHisProIleSer 360
DB 1580 CCTGTGCCAGCTCTGCTACCTCAGAGGACAGCAAGGCGCGGATCCGCGCAGCG 1639
QY 360 alProSerValProSerValGlnProSerIleuLeuIysAspSerProIleuValGlnAla 380
DB 1640 TGCCAGCGTGGCCAGCGTGCAGCCGCTGCTGTGAAGACAGCCCGCTGGTACAGGCTC 1699
QY 380 euSerGlyLeuHisIleuSerAlaAspArgLeuGluIysAspSerAsnIysLeuIysArgSer 400
DB 1700 TCAGTGGGCTGCACCTGTCGAGCAGAGCTGGAAACAGCAATTAAGCTCAAGCTTCT 1759
QY 400 heSerLeuAspIleIysSerValSerTyrSerIleAsnThrAlaAsnIleAsnIleGlyP 420
DB 1760 TCTCTGGAATCAATCAATGATTCATATTCAGCCAGCTGACAGATCTTACATGCT 1819
QY 420 heSerSerGluAspAlaLeuGluTyrTyrIlePheProSerThrThrIleuAspGlyTyr 440
DB 1820 TCTCTCATCAGAAAGTGTGTTGGAATCTTCAAACTTCCACTCTCTGAGTGGAGCA 1879
QY 440 snIysLeuCysGlnPheSerProValGlnIleuSerGluIlnThrProGluThrSerP 460
DB 1880 ACAAGCTATGCAAGTCTCTCCCTGTTCAAGAACTATCGAGCAGACTCCGGAACAGT 1939
QY 460 roAspIysGluIysAlaSerIleProIysIysLeuGlnThrAlaArgProSerArgSer 480
DB 1940 CTGATAGAGAGAGCAGATCCCAAGAGCTCAGACGCGCGGCTTCAAGCACCC 1999
QY 480 InSerIysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500

```

Db      2000 AGAGCAGAGGATTCGATTCGGTCAGAAACGACGACGATGGCACCCGCCAGAGGTCCTTT 2059
Qy      500 euserProleuHi:earSerGlySerValGluAapAntyRhi:sthrSerPheLeuPheG 520
Db      2060 TATCTCACCTGACATCCAAAGTGGAGCGTGAGACATATTCACACACGCTTCTTTTCG 2119
Qy      520 lyeuserHrSerGlnGlnHisLeuThrIysSerIagIlyLeuGlyLeuIyGlyTyrH 540
Db      2120 GCGTTTCCACCCACGACGACGACCTCACGAAGTCTGCGCTGGCGCTTAAAGGCGCTGGC 2179
Qy      540 iserAapIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTriPtyPheA 560
Db      2180 ACTCGGATATTCGGCCCCCACCACCTCTACCCCTTCCCTGACACGACGCTGATTTTCG 2239
Qy      560 lathrGluSerSerH:apheTySerAlaSerAlaIleTyGlyGlySerAlaSerTyS 580
Db      2240 CCACAGAGTCTCACCTTCTACTGCTGCTCAGCCCATCTACGAGGACAGTCCAGTTACT 2299
Qy      580 eprIatYrSerCySerGlnLeuProThrCySerGlyAapGlnValTySerValArgArgA 600
Db      2300 CTGCTTACAGCTGACGACGACGCTCCCACTTGCGAAGCAAGTCTATCTGTCGCGAGGC 2359
Qy      600 rglInlySProSerAapArgAlaAapSerArgArgSerTriPti:sgIuGluSerProPheG 620
Db      2360 GCGAGAAAGCCAACTGACAGAGCTGACTCGCGGCGGAGCTGGCATGAAGAGGCCCTTTG 2419
Qy      620 luleGlnPheIyArGArgSerCySerGlnMetGluPheGlyGluSerIleMetSerGluA 640
Db      2420 AAAAGCAGTTTAAACGACAGAGCTGCCAAATGGAAATTGGAGAGACATCATGTGACAGAG 2479
Qy      640 aapArgSerArgGluGluLeuGlyIyValGlySerGlnSerSerPheSerGlySerMetG 660
Db      2480 ACAGGTCACGGGAAAGCTGGGGAAGTGGCGACGTACGTTTTCGGGCGACGATGG 2539
Qy      660 luleIleGluValSer 665
Db      2540 AATCATTTGAGCTCTCC 2556

RESULT 9
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589)...(12583)
US-09-816-494-1

Alignment Scores:
Pred. No.: 0 Length: 3544
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 9 Gaps: 0

US-10-029-345a-109 (1-665) x US-09-816-494-1 (1-3544)

```

```

Qy      1 MetAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db      589 ATGGCCCATGAGTATGATTTGAACTCAAAATTTTACTGAGAGGTTGGCTTCCTGGAA 648
Qy      21 SerGlyThrGluIyValLeuLeuIleAapSerArgProPheValGluTyranThrSer 40
Db      649 AGTGAAGAGGAAAAAGTGTGCTGAATATGATGAGCGGCAATTTGTGAATATCAATATAC 708
Qy      41 HisIleLeuGlnAlaIleAenIleAenCySerIyLeuMetIyAArgArgLeuGln 60
Db      709 CACATTTTGAAGCCCTTATATCACTCTCCAACTTATGAAAGGTTGCAACAG 768
Qy      61 AapIyValLeuIleHrGluLeuIleGlnHisSerAlaIyHisIyValAapIleAap 80
Db      769 GACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGCAAAACATTAAGTTGACATTGAT 828
Qy      81 CySerGlnIyValValValTyranPginSerSerGlnAapValAlaSerLeuSerSer 100
Db      829 TGCAGTCAGAGGTTGTATTTACATCAAGTCCCAAGATGTTCCCTCTCTTCA 888
Qy      101 AapCyPheLeuThrValLeuLeuGlyIyIyLeuGluIySerPheAanSerValHisLeu 120
Db      889 GACTGTTTCTCAGCTGTACTTCTGGGTTAACTGAGAAAGGTTCAACTCTGTTCACTTG 948
Qy      121 LeuAlaGlyIyPheAlaGluPheSerArgCyPheProGlyLeuCySerGluGlyIySer 140
Db      949 CTTCAGAGTGGGTTTGTAGATTCCTCGTGTTCCTGCGCTCTGTGAAGAAATATCC 1008
Qy      141 ThrLeuValProThrCySerIleSerGlnProCySerLeuProValAlaAenIleGlyProThr 160
Db      1009 ACTCTAGTCCCTACCTGACATTTCTCAGCTTCTTACCTGTGCCAACTTGGGCCAAC 1068
Qy      161 ArgIleLeuProAenLeuTyranLeuGlyCySerIyAArgAapValLeuAenIyGluLeu 180
Db      1069 CGAATTCCTCCAACTTTATCTGTGCTCCAGCGAGTGTCTTCAACAGAGCTGAT- 1127
Qy      181 -GlnGlnAenGlyIleGlyTyranValLeuAenAlaSerTyranCyPProIyProAap 200
Db      1128 GCGAGCAATGGAGTGTATGTATGTTAATGCCAGCAA-TACCTGTCCAAAGCTGTAGCT 1186
Qy      200 heIleProGluSerHisPheLeuArgValProValAanAapSerPheCySerGluIyIle 220
Db      1187 TTTATCCCGCAGTCTCATTTCTCGCGTGTCTGTGAATGACAGCTTTGTGAGAAATTT 1246
Qy      220 eupProTriPheAapIySerValAapPheIleGluIyAlaIyAlaSerAenGlyCyArg 240
Db      1247 TCCCGGTGTGGAACAATAGTATGATTTCAATGAGAAACAAGCCCTCAATGATGATG 1306
Qy      240 AlLeuValHisCySerLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyran 260
Db      1307 TTTCTAGTGACGTGTTAGCTGGAGATCTCCGCTCCGCCACATCGCTATCGCTATACAT 1366
Qy      260 eIyAArgMetAapMetSerLeuAapGluAlaTyranPheValIyGluIyArgProT 280
Db      1367 TGAAGAGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAAGAAAAAGCCCTA 1426
Qy      280 hrIleSerProAanPheAenPheLeuGlyGlnLeuLeuAapTyranIyGluIyValIyA 300
Db      1427 CTATATCTCCAAACTTCAATTTTTCGGGCAACTCTGAGCATATGAGAAAGATTAAGA 1486
Qy      300 sngIlnThrGlyAlaSerGlyProIySerIyLeuIyLeuLeuHisIleLeuGluIyValP 320
Db      1487 ACCAGACTGAGACATCAGGCGCAAGAGCAAACTCAAGTGTGTGACCTGAGAAACC 1546
Qy      320 sngIlyProValProAlaValSerGluGlyGlyIyIySerGluTyranProAapSerProP 340
Db      1547 ATGAACCTGTCCCTGCTGTCTCAGAGGCTGGACAAAGAACAGAGCCCTCTCAGTCCAC 1606
Qy      340 roCyAlaAapSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaSerV 360
Db      1607 CCTGTGCGGACTGTGCTACCTCAGAGGCGAGCAAAAGCCCGTGATCCCGCAGCG 1666
Qy      360 aIProSerValProSerValGlnProSerLeuLeuGluAapSerProLeuValGlnAlaL 380

```



```
Db 1667 TGCCAGCGTCCGCGCGTGCAGCGCGTGCCTGTTAGAGACAGCCCGCTGTACAGGCC 1726
Qy 380 euserGlyLeuHiLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerp 400
Db 1727 TCAGGGGCTCAGCTGTCTCCGACAGAGCTGGAGAACAGCAATTAAGCTCAAGGCTTCT 1786
Qy 400 heserLeuAspIleLysSerValSerTyrSerAlaSerMetAlaIleSerLeuHISglYp 420
Db 1787 TCTCTGGATATCAATCACTTTCATTTTCAGCCAGATGGAGCATCTTACATGGCT 1846
Qy 420 heserSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThra 440
Db 1847 TCTCTCATCAGAAAGATGCTTTGGAAATCAAAACCTTCCACTACTCTGGATGGAGCA 1906
Qy 440 snLysLeuCySGlnPheserProValGlnGluLeuSerGluGlnThrProGlnThrSerp 460
Db 1907 ACAAGATATGCCAATTCCTCCCTGTTCAGGAATCATGGAGCAGACTCCGAAACCAAGTC 1966
Qy 460 roAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerp 480
Db 1967 CTGATTAAGAGAAAGCAGCATCTCCCAAGAGCTGCAGACCGCCAGGCTTTCAGACAGCC 2026
Qy 480 InserLysArgLeuHiSerValArgThrSerSerGlyThrAlaGlnArgSerLeu 500
Db 2027 AGAGCAAGCAGATTCGATTCGTCAGAACCAAGCAGAGTGGCAGCCCGCCAGAGGTCCTTT 2086
Qy 500 euserProLeuHiSArgSerGlySerValGluAspAsnTyrHISThrSerpPheLeuPhag 520
Db 2087 TATCTCCACTGCATCGAAGTGGAGGCGTGGAGGACATTTACACACAGAGTTCCTTTTG 2146
Qy 520 LysSerThrSerGlnGlnHISLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrPh 540
Db 2147 GCCTTTCCACAGCCAGCAGCACTCAGCAAGTCTGCTGGCTGGCTTAAAGGCTGGCC 2206
Qy 540 isserAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerpTyrPheA 560
Db 2207 ACTGGATATCTTGGCCCCCAGACCTTAACCTTCCCTGACAGCAGACTGGTATTTTG 2266
Qy 560 laThrGluSerSerHISpHeTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrS 580
Db 2267 CCACAGAGTCTCAGACTTCTACTCTGCTCAGCCATCTACGAGGCACTGGCCAGTTACT 2326
Qy 580 exLaIyTyrSerCySGerGlnLeuProThrCySGIYAspGlnValTyrSerValArgArg 600
Db 2327 CTGCCCTAAGCTGCAGCCAGCTGCCACTTGGCGAAGCAGCATATCTGTGGCCAGGC 2386
Qy 600 rGlnLysProSerAspArgAlaAspSerArgArgSerTyrHISgluGluSerProPhag 620
Db 2387 GCGAGAACCAAGTGCAGAGCTGACTCGCGGGGAGCTGGCATGAAGAGACCCCTTTG 2446
Qy 620 LuLysGlnPheLysArgArgSerCySGInMetGluPheGlyLysSerIleMetSerGlu 640
Db 2447 AAAAGCAGTTTAAACGACAGAGCTGCCAAATGGATTGGAGAGCATCATGTACAGGA 2506
Qy 640 snArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySermetG 660
Db 2507 ACAGGTCACGGGAAGAGCTGGGGAAGTGGCAGTCACTTTCCTTTCGGGAGCATGG 2566
Qy 660 LuIleIleGluValSer 665
Db 2567 AAATCATTAAGGCTCTCC 2583
```

RESULT 10
US-10-377-072-25

```
/ Sequence 25, Application US/10377072
/ Publication No. US20040009501A1
/ GENERAL INFORMATION:
/ APPLICANT: Millennium Pharmaceuticals Inc.
/ APPLICANT: Curtis, Kory A.J.
/ APPLICANT: Logan, Thomas Joseph
/ APPLICANT: Glucksmann, Maria A.
/ APPLICANT: Meyers, Rachel E.
```

```
/ APPLICANT: Williamson, Mark J.
/ APPLICANT: Rudolph-Owen, Laura A.
/ APPLICANT: Chun, Miyoung
/ APPLICANT: Teal, Feng-Ying
/ TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,
/ TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES
/ TITLE OF INVENTION: AND USES THEREFOR
/ FILE REFERENCE: MP103-018QNMIM
/ CURRENT APPLICATION NUMBER: US/10/377,072
/ PRIOR FILING DATE: 2003-02-27
/ PRIOR APPLICATION NUMBER: US 09/895,860
/ PRIOR FILING DATE: 2001-06-29
/ PRIOR APPLICATION NUMBER: US 60/215,370
/ PRIOR FILING DATE: 2000-06-29
/ PRIOR APPLICATION NUMBER: US 09/723,806
/ PRIOR FILING DATE: 2000-11-28
/ PRIOR APPLICATION NUMBER: US 60/187,455
/ PRIOR FILING DATE: 2000-03-07
/ PRIOR APPLICATION NUMBER: US 09/843,297
/ PRIOR FILING DATE: 2001-04-25
/ PRIOR APPLICATION NUMBER: US 60/199,801
/ PRIOR FILING DATE: 2000-04-26
/ PRIOR APPLICATION NUMBER: US 09/861,801
/ PRIOR FILING DATE: 2001-05-21
/ PRIOR APPLICATION NUMBER: US 60/205,508
/ PRIOR FILING DATE: 2000-05-19
/ PRIOR APPLICATION NUMBER: US 09/816,494
/ PRIOR FILING DATE: 2001-03-23
/ PRIOR APPLICATION NUMBER: US 09/815,419
/ PRIOR FILING DATE: 2001-03-22
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 25
/ LENGTH: 3544
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (589)...(2586)
/ US-10-377-072-25

Alignment Scores:
Pred. No.: 0 Length: 3544
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: Gaps: 0

US-10-029-345A-109 (1-665) x US-10-377-072-25 (1-3544)
Qy 1 MetAlaHISgluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 589 ATGGCCCATGAGATGATTTGGAACTCAAAATTTGTTACTGAGAGGTTGGCTGCTGCGAA 648
Qy 21 SerGlyThrGluLysValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40
Db 649 AGTGAACGGAAAGAGTGGCTGTAATTTGATAGCGGCAATTTGTGGAAATCAATCATCC 708
Qy 41 HIsIleLeuGluAlaIleAsnIleAsnCySGerLysLeuMetLysArgArgLeuGln 60
Db 709 CACATTTTGGAGCCATTATATCACTCTCCAACTTATGAAGCGAAGTTGCACACG 768
Qy 61 AspLysValLeuIleThrGluLeuIleGlnHISerAlaLysHISLysValAspIleAsp 80
Db 769 GACAAAGTTTATTTACAGAGCTCATCCAGCATTCAGCAACATTAAGTTGACATTGAT 828
Qy 81 CySGerGlnLysValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
Db 829 TGCAGTCAGAAAGGTTTATTTACATCAATCAAACTCCCAAGAGTGGCTTCTCTCTTCA 888
Qy 101 AspCySPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHISLeu 120
```



```

Db      889  GACTGTTTCTGACTGCTGCTGGGTTAACTGGAGAAAGCTTCAACTCTGTTTCACTCG 948
Qy      121  LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuGlyGluGlySer 140
Db      949  CTTGCGAGGGGGTTCCTGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008
Qy      141  ThrLeuValProThrCysAlaIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
Db      1009  ACTTATGCTCCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
Qy      161  ArgIleLeuProAsnLeuTyrLeuGlyCysGlnArgAspValLeuAsnIleGlyLeuIle 180
Db      1069  CGAATTCCTCCCAATCTTTATCTTGGCTGCGACGAGGAGATGCTCTCAACAGAGCTGATG 1127
Qy      181  -GlnGlnGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 200
Db      1128  GCAGCAGAAATGGATGGTATGCTGTTAAATGCCAGCAA-TACCTGTCAGAAAGCTGACT 1186
Qy      200  heIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluValIleL 220
Db      1187  TTATCCCGCAGTCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246
Qy      220  euProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAspSerAsnGlyCysV 240
Db      1247  TCCCGTGGTTGGACAAATCAGTAGATTTTCATTGAGAAACAAAGCCTCCATGATGATG 1306
Qy      240  alleuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleM 260
Db      1307  TTCTAGTGACATGTTAGCTGGGATCTCCGCTCCGCCACCATCGCTATCGCTATCATCA 1366
Qy      260  eLysArgMetSerPheMetSerLeuAspGluAlaIleTrpArgPheValLysGluLysArgPro 280
Db      1367  TGAAGGAGATGGACATGCTTTAGATGAAGCTTACGATTGGAAAGAAAAAGCCCA 1426
Qy      280  hTrIleSerProAsnPheAsnPheLeuGlyGlnLeuAspTyrGluLysIleLysA 300
Db      1427  CTRATCTCCAAACTTCATTTCTGGGCCAACTCCTGGAATGAGAAAGATTAA 1486
Qy      300  bngIlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLysLeuLysLeuLysPro 320
Db      1487  ACCAGACTGGACATCAGGGGCCAAAGACAACTCAAGCTGCTGACCTGGAGAGCCAA 1546
Qy      320  bngIlnProValProAlaValSerGluGlyGlnLysSerGlnTrpProLeuSerProp 340
Db      1547  ATGAACCTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1606
Qy      340  roCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerV 360
Db      1607  CCTGTGCCACTCTGCTTACTCTAGAGGCGACGAGACAAAGCCCGCTGATCCCGCCAGCG 1666
Qy      360  aIProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAlaL 380
Db      1667  TCCCGCAGCTGCCACGCTGAGCGCTGCTGTTAAGACAGCCCGCTGTTACAGGGCC 1726
Qy      380  euSerGlyLeuHisIleuSerAlaAspArgLeuGlnAspSerAsnLysLeuLysArgSerP 400
Db      1727  TCAGTGGGCTGACCTGCTCGCAGACAGAGCTGGAAACAGCATTAAGCTCAAGCGTTCT 1786
Qy      400  heSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyP 420
Db      1787  TCTCTCTGATATCAATCAATGATTTCTATATTCAGCCGCAATGGCAGCATCTTCAATGGCT 1846
Qy      420  heSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThrA 440
Db      1847  TCTCCTCATCAGAAAGATGCTTTGGAATATCAAAACCTTCCACTACTCGATGGAGCA 1906
Qy      440  snLysLeuGlyGlnPheSerProValGlnGluLeuSerGluGlnTrpProGluTrpSerP 460
Db      1907  ACAAGCTAATGCGAGTCTCCCTGTTACAGAACTATCGAGACAGATCTCCGAAACCAAGTC 1966
Qy      460  roAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerG 480

```

```

Db      1967  CTGATPAGGAGGAGCCAGCATCCCAAGAGCTGACAGCCGCGACGCTTCAGACAGCC 2026
Qy      480  InSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuL 500
Db      2027  AGAGCAAGGATTTGATTCGTGTCAGAACGAGCAGCTGGACCCGCCAGAGGCTCCTTT 2086
Qy      500  euSerProLeuHisArgSerGlySerValGluAspAsnTrpHisThrSerPheLeuPheG 520
Db      2087  TATCTCCTCATTCGATGAGTGGAGGAGCTGGAGACAAATTAACACACAGCTTCTTTTG 2146
Qy      520  LysSerThrSerGlnGlnHisIleuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrP 540
Db      2147  GCCTTTCACCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2206
Qy      540  iSerAspIleLeuAlaProGlnTrpSerThrProSerLeuThrSerSerTrpTyrPheA 560
Db      2207  ACTCGGATATCTTGGCCCCCAGACCTTACCTTCCCTGACAGCAGCAGCAGCAGCAGCAG 2266
Qy      560  LaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrS 580
Db      2267  CCACAGAGTCTTCACTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2326
Qy      580  eAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgA 600
Db      2327  CTGCTCAGAGCTGACAGCCAGCTGCCACTTGGGAGAGCAAGTCTATTTCTGTGCGCAGG 2386
Qy      600  rGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSerProPheG 620
Db      2387  GGCAGAGCAGCAAGTACAGAGCTGACTCGCGCGAGCTGGGATGAGAGAGAGAGAGAG 2446
Qy      620  LuLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMetSerGluA 640
Db      2447  AAAAGAGTTTAAAGCGCAAGCTGCTCAATGAAATTTGAGAGACATCATGTCTGAGA 2506
Qy      640  snArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetG 660
Db      2507  ACAGGTCAAGGGAAGAGCTGGGAAAGTGGGAGTCAAGTCAAGTCTTTTGGGCGAGATGG 2566
Qy      660  LuIleIleGluValSer 665
Db      2567  AATCATTCAGGCTCTCC 2583

RESULT 11
US-10-425-114-26234
; Sequence 26234, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26234
; LENGTH: 3625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-028-H6_FLI
US-10-425-114-26234

Alignment Scores:
Pred. No.: 0
Score: 472.00
Percent Similarity: 99.40%
Best Local Similarity: 99.40%
Query Match: 70.98%
Length: 3625
Matches: 663
Conservative: 0
Mismatches: 2
Indels: 4

```

DB: 13 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-425-114-26234 (1-3625)

QY 1 MetLahsGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20

DB 692 ATGGCCCATGAGATGATTGGAACTCAATTTGTACTGAGAGTGGTGGCTGCTGGA 751

QY 21 SerGlyThrGluValLeuLeuIleAspSerArgProPheValGluTyrAsnThrSer 40

DB 752 AGTGAACGAGAAAGAGCTGCTCAATGATAGCCGCGCATTTGGAAATCAATACATCC 811

QY 41 HisIleLeuGluAlaIleAsnIleAsnCySerLysLeuMetLysArgLeuGln 60

DB 812 CACATTTTGGAAAGCCATTAATATCACTGCTCCAGCTTATGAAAGCAAGTTGCCAAG 871

QY 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAspIleAsp 80

DB 872 GACAAAGTGTAAATTAACAGAGCTCATCAGCATTCACGAAACATAAAGTTGACATTTGAT 931

QY 81 CySerGlnLysValValIleTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100

DB 932 TGCACTAGAAAGGTGATGATTACAGATCAAAAGCTCCACAGATGTGCTCTCTCTTCA 991

QY 101 AspCyPheLeuThrValIleLeuGluLysLeuGluLysSerPheAsnSerValHisLeu 120

DB 992 GACTGTTTCTCACTGACTTCTGGGTAACGTGAGAAAGCTTCAACTCTGTTCACTG 1051

QY 121 LeuAlaGluGlyPheAlaGluPheSerArgCyPheProGlyLeuCyGluGlyLysSer 140

DB 1052 CTTGCAAGTGGGTGTGTGAGTCTCTGTTGTTTCCCTGGCCTCGTAAAGGAAATTC 1111

QY 141 ThrLeuValProThrCysIleSerGlnProCyLeuProValAlaAsnIleGlyProThr 160

DB 1112 ACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTTGCCAACTGGGCCAAC 1171

QY 161 ArgIleLeuProAsnLeuTyrLeuGlyCySGlnArgAspValLeuAsnLysGluLeuIle 180

DB 1172 CGAATCTTCCCAATCTTATCTTGGCTCCAGCGAGATGTCCTCAACAGAGCTGAT 1230

QY 181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCyProLysProAsp 200

DB 1231 GCGAGAGATGGATGGATTGATGTGTTAAATGCGAGCAAA-TACCTGTCCAAAGCCTGACT 1289

QY 200 IleLeuProLysSerHisPheLeuArgValProValAsnAspSerPheCyGluLysIle 220

DB 1290 TTATCCCGAGTCTCATTTCCGCGGTGCTCGTGAATACAGCTTTGTGGAATAATTT 1349

QY 220 LeuProLysSerLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCySV 240

DB 1350 TGCCTGGTGGACAAATCAGATGATTTCAATTGAGAAAGCAAAAGCCTCAATGATGTG 1409

QY 240 AlaLeuValHisCySerLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 260

DB 1410 TTCTAGTGAAGCTTTAGCTGGATCTCCCGCTCCGCAACCATTCGCTTACATCA 1469

QY 260 eLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgProT 280

DB 1470 TGAAGAGGATGACATGCTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTA 1529

QY 280 HisIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIleLysAla 300

DB 1530 CTATATCTCCAAACTTCAATTTTCTGGGCGCAACTCTGAGCTATGAGAAAGATTTAGA 1589

QY 300 snGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisIleGluLysProAla 320

DB 1590 ACCAGACTGGAGCATCAAGGCCAAAGAGCAAACTCAAGCTCTGCACTGAGAAAGCCAA 1649

QY 320 snGluProValProAlaValSerGluGlyGlyLysSerGluThrProLeuSerProp 340

DB 1650 ATGAACCTGCTCCCTGCTCTCAGAGGGGTGACAGAAAGGAGACGCCCTCTGCTCAC 1709

QY 340 roCysAlaAspSerAlaThrSerGluAlaIleGlyGlnArgProValHisProAlaServ 360

DB 1710 CCTGTGCCACTCTCTACTCTCAGAGCAGCAGCAAAAGGCCCTGATCCCGCAGGC 1769

QY 360 aLProSerValProSerValGlnProSerLeuLeuGlnAspSerProLeuValGlnAla 380

DB 1770 TGCCTCAGCTGCGCCAGCTGACAGCCGTGCTGTTAGAGACAGCCGTGGTACAGGGCC 1829

QY 380 euSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerP 400

DB 1830 TCAGTGGGCTGCACTGCTCCGACAGCAGCTGGAAGACAGCAATAGCTCAAGCTTCT 1889

QY 400 heSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyP 420

DB 1890 TCTCTGTGAATACAAATCAATTTCAATTCAGCCAGCATGCGACATCTTACATGAGCT 1949

QY 420 heSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThr 440

DB 1950 TCTCTCATCAGAAAGATCTTTGAAATCTCAAACTTCCACTACTCTGATGGAGCA 2009

QY 440 snLysLeuCyGlnPheSerProValGlnLysSerGluGlnThrProGluThrSerP 460

DB 2010 ACAAGCTATGCGCAGTCTCCCTGTTCAAGAACTATCGAGACAGATCCCGAAACAGTCC 2069

QY 460 roAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerG 480

DB 2070 CTGATTAAGAGAAAGCCAGCATTCCTCCAGAAAGCTGACAGCCCGCAGCTTTCAGACAGCC 2129

QY 480 InSerLysArgLeuHisSerValArgThrSerSerSerSerGlyThrAlaGlnArgSerLeu 500

DB 2130 AGAGCAAGCATTCATTCGCTGACAGAACAGCAGCAGATGACCCGCAAGGTCTCTTT 2189

QY 500 euSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPheLeuPheG 520

DB 2190 TATCTCAGCTGACATGAAAGTGGAGCGTGGAGGACAAATTACACACAGCTTCTTTTCG 2249

QY 520 LysSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLysGlyTyrHis 540

DB 2250 GCCTTTCACACAGCAGCAGCAGCAGCTTCCAGAAAGTGTGCTGGGCTTTAAAGGCTGGC 2309

QY 540 IsSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheAs 560

DB 2310 ACTCGATATCTTGGCCCCCAGACCTCTACCTCCCTTCCGACACAGCTGGTATTTTG 2369

QY 560 laThrGlnSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrS 580

DB 2370 CCAGAGTCTTCACTTCTACTCTGCTCAGCAGCTTCAACGAGGACAGTGCAGTACT 2429

QY 580 eAlaTyrSerCySerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600

DB 2430 CTGCTTACAGCTCAGCAGCAGCTGCCCATTTGCGGAGACCAAGTCTATCTGTGCGCAGGC 2489

QY 600 rgGlnLysProSerAspArgAlaAspSerArgArgSerTyrHisGluGlnSerProPheG 620

DB 2490 GGCAGAAAGCCAAATGACAGAGCTGACTCGGGGGAGCTGGCATGAAGAGAGCCCTTTG 2549

QY 620 LysGlnPheLysArgArgSerCyGlnMetGlnPheGlyGlnSerIleMetSerGluAla 640

DB 2550 AAAAGAGTTTAAACGACGAGAGCTGCAATATGAAATTTGGAAGAGCAATCAATGAGAA 2609

QY 640 snArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetG 660

DB 2610 ACAGTCAAGGAGAAAGCTGGGGAAAGTGGGAGTCAAGTCAAGCTTTTGGGACAGATGG 2669

QY 660 LulleIleGluValSer 665

DB 2670 AAATCAATTGAGGTCTCC 2686

RESULT 12

US-10-343-357-17

; Sequence 17, Application US/10343357

; Publication No. US20040058341A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom

APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Jayalaxmi
APPLICANT: YAO, Monique G.; BUREFORD, Neil
APPLICANT: WANG, Yumei E.; STEWART, Elizabeth A.
APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.
APPLICANT: LEE, Ernestine A.; HARALJA, April J.A.
APPLICANT: LU, Dyung Aina M.; TRIBOULEY, Catherine M.
APPLICANT: GRIFFIN, Jennifer A.; BAUGHN, Mariah R.
APPLICANT: YUE, Henry; WARREN, Bridget A.
APPLICANT: NGUYEN, Daniel B.; CHAWLA, Narinder K.
APPLICANT: KEARNEY, Liam
TITLE OF INVENTION: PROTEIN PHOSPHATASES
FILE REFERENCE: PI-0173 PCT
CURRENT APPLICATION NUMBER: US/10/343,357
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: PCT/US01/23716
PRIOR FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,679
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/223,272
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/224,309
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US 60/226,728
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 60/229,254
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: US 60/231,366
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 17
LENGTH: 3766
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040058341A1 7480570CBI
US-10-343-357-17

Alignment Scores:
Pred. No.: 0 Length: 3766
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 13 Gaps: 0

US-10-029-345a-109 (1-665) x US-10-343-357-17 (1-3766)

QY 1 MetAlAHlAGlUeMellleglYThrGlnlIleValThGluArgleuValAlaleuLeuGlu 20
DB 538 ATGGCCCATGAGATGATGGAGACTCAAAATTGTAATCGAGAGGTGGCTGCTGGCGAA 597
QY 21 SerG1YThrGluYerValleuLeuIleApsSerArgProPheValGluTYrAanThSer 40
DB 598 AGTGGAAACGGAAGAGTGCCTGCTAATTGATGACCGGCATTGTGGAAATCAATACATCC 657
QY 41 HsIlleuGluAlaleuAnlleuApsSerIleuLeuIleValArgArgleuGln 60
DB 658 CACATTTTGGAGCCATTAATATCACTGCTCCAGCTTATGAAGCGAGTTGCAACG 717
QY 61 AspIleValleuIleThrGluLeuIleGlnHsSerAlaYshIleValIleAps 80
DB 718 GACAAAGTGTAAATTCAGAGCTCATCCAGCATTCAGCGAAACATAGAGTTGACATTGAT 777
QY 81 CysSerGlnIleValIleValTYrApsGlnSerSerGlnApsValAlaSerIleuSer 100
DB 778 TGCAGTCAGAGAGGTGAGTTAGATCAAGCTCCCAAGATGGCTTCTCTCTTCA 837
QY 101 AspCysPheLeuThrValleuLeuGluYlyArgleuGluYserPheAnSerValIleu 120
DB 838 GACGTGTTTCTCACTGTAATTCTGGGTAACTGAGAAAGACTTCAACTGTTCACTG 897

QY 121 LeuAlaGlyIlePheAlaGluPheSerArgCysPheProGlyLeuGlySer 140
DB 898 CTTCGACAGTGGGTTTGTGAGTTCTCTCTGTTCCCTGGCCTCTGTGAAGAAATTC 957
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAnlleGlyProThr 160
DB 958 ACTCTAGTCCCTACCTGACATTTCTCAGCTTCTCTGTTGCCAACTGGGCCAACCC 1017
QY 161 ArgIleLeuProAnlleuTYrIleuGlyCysGlnArgAspValleuAnlleuIle 180
DB 1018 CCAATTCCTCCAAATCTTATCTTGGCTGCCAGCGAGATGCTTCAACAGAGCTGAT- 1076
QY 181 -GlnGlnAengIyIleGlyTYrValleuAnlleuAserTYr-ThrCysProIysProAps 200
DB 1077 GAGCAGAGTGGGATTTGTTATGTGTTAAATGCCAGCA-TACCTGTCCAAAGCTTGACT 1135
QY 200 HsIlleProGluSerHsIlePheLeuArgValProValAnlleuApsSerPheCysGluYle 220
DB 1136 TTATCCCGAGTCTCATTTCTGCGGTGCTGTGATGACAGCTTTGTGAGAAATTT 1195
QY 220 euProTTrleuApsIleSerValIlePheIleGluYalAlaYshIleAserAnlleCys 240
DB 1196 TCCCGGTGTGGACAAATCAGATGATTTGATGAGAAACCAAGCTTCAATGATGTG 1255
QY 240 allleuValHsCysIleuAlaGlyIleSerArgSerAlaThrIleAlaIleATyrIle 260
DB 1256 TTCTAGTCAGCTGTTTACGTGGGATTCCTCGCTCCGCACTGCTTAATGCTTAATCA 1315
QY 260 eClyeArgMetApsMetSerLeuApsGluAlaTYrArgPheValYlyGluYArgPro 280
DB 1316 TGAAGAGATGGACATGCTTTTAGATGAGCTTACAGATTTGGAAAGAAAGACCTTA 1375
QY 280 HsIlleSerProAnlleuPheAnlleuGlyGlnleuLeuApsTYrGluYlyIleYal 300
DB 1376 CTATATCTCCAACTTCAATTTTCTGGGCCCACTCTGACATGAGAAAGATTTAAG 1435
QY 300 sngIlnThrGlyIleAserGlyProIysSerIleuYserLeuHsIleuGluYPro 320
DB 1436 ACCAGCTGAGATAGAGGCCCAAGAGCAAACTCAAGCTGCTGCACTGGAGAACCA 1495
QY 320 sngIlnProValProAlaValIserGluGlyIleYlySerGluThrProLeuSerPro 340
DB 1496 ATGAACCTTCCCTGCTGCTGCTCAGAGGTGAGACAGAAAGCGAGAGCCCTCAGTCC 1555
QY 340 rOCyAlaApsSerAlaThrSerGluAlaIleGlyGlnArgProValIleProAlaSer 360
DB 1556 CCTGTGCCACTGCTGCTACCTCAGAGCGACAGCAAGAGCCCTGCAATCCGCCAGCG 1615
QY 360 alProSerValProSerValGlnProSerIleuGluApsSerProLeuValGlnAla 380
DB 1616 TGCACAGCTGCCCGCAGCGTGCAGCGCTGCTGTTAGAGACAGCCGCTGGTACAGGCG 1675
QY 380 euSerGlyLeuHsIleuSerAlaApsArgleuGluApsSerAnlleuYlyArgSer 400
DB 1676 TCAAGTGGCTGACACCTGTCGCGAGACAGGCTGGAACAGCAATAGCTCAAGCTTCC 1735
QY 400 HsSerLeuApsIleIleYserValSerTYrSerAlaSerMetAlaIleSerLeuHsGly 420
DB 1736 TCTCTCTGATATCAAAATCAGTTTCAATTTACGCGCATGCGAGATCTTACATGGCT 1795
QY 420 HsSerSerGluApsAlaIleuGluTYrTYrIleYArgSerPheThrIleuApsGlyTh 440
DB 1796 TCTCTCATCAGAAAGTGTGGAAATACATCAAACTTCCACTACTCTGGATGGACCA 1855
QY 440 snIleuLeuCysGlnPheSerProValGlnGluSerGluGlnThrProGluThrSer 460
DB 1856 ACAAGCTTAGCCAGTCTCCCTGTTCAAGAACTATCGAGAGAGACTCCGAAACCAAGT 1915
QY 460 rOCAsIleGluGluAlaSerIleProYlyIleuGluIlnThrAlaArgProSerApsSer 480
DB 1916 CTGATTAAGAGAGAGCCAGCATCTCCCAAGAGCTGACAGCCGCTTCAAGACAGCC 1975
QY 480 IlnSerIleArgleuHsIleSerValArgThrSerSerSerGlyThrAlaIlnArgSerLeu 500

DB 1976 AGAGCAAGCATTCGATTCGGTCAGAACAGAGAGAGTGGACCCCGCAGAGGTCCCTTT 2035
QY 500 euseProleuHiaSergGlySerValGluAspSerTrpHisThrSerPheLeuPheG 520
DB 2036 TATCTCCACTGATCGAAGTGGAGCCGTGGAGACAAATTACACACAGCTTCTTTTCG 2095
QY 520 lYleuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTrpH 540
DB 2096 GCCTTTCACACGACGACGACGACCTTCAGCAAGTCTGCTGAGCTTGGGCTTTAAGGGCTGGC 2155
QY 540 lSerAspGlnLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpLysPheA 560
DB 2156 ACTGGATATCTTGGCCCCCAGACCTTCACCTCCCTCCAGACAGAGCGGTGATTTTG 2215
QY 560 lATHrGluSerSerHisPheThrSerLysSerAlaIleTrpGlyGlySerAlaSerTrpS 580
DB 2216 CCACAGAGCTCTTCACACTTCTACTCTGCTGACCATCTACGAGGACAGTCCAGTTACT 2275
QY 580 eAlaTrpSerCysSerGlnLeuProThrCysGlyAspGlnValTrpSerValArgArgA 600
DB 2276 CTGCTTCAGCTGACGACGAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCGAGGC 2335
QY 600 rGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSerProPheG 620
DB 2336 GGCAGAAAGCCAGAGACAGAGCTGACTCGCGCGAGAGCTGCATGAAGAGAGCCCTTGG 2395
QY 620 lYlYsGlnPheLysArgArgSerCysGlnMetGlnPheGlyGlyGlySerLysMetSerGlnA 640
DB 2396 AAAAGCATTTAAACGACGAAAGCTGCCAATTCGAAATTTGAGAGAGCATCATGTCAAGA 2455
QY 640 snArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetG 660
DB 2456 ACAGGTACCGGAGAAAGCTGGGAGAAAGTGGGAGTCACTTACTTTTGGGAGAGATGG 2515
QY 660 lYlLeuLeuGluValSer 665
DB 2516 AAATCATTTGAGGTCTCC 2532
RESULT 13
US-10-648-593-115
; Sequence 115, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 115
; LENGTH: 4790
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-648-593-115
Alignment Scores:
Pred. No.: 0 Length: 4790
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 17 Gaps: 0
US-10-029-345A-109 (1-665) x US-10-648-593-115 (1-4790)
QY 1 MetalAlaHisGluMetIleGlyThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
|||||

DB 184 ATGGCCCATGAGATGATTGAACTCAAAATTGTTACTGAGAGGTGGCTGCTGCGAA 243
QY 21 SerGlyThrGluLysValLeuLeuLysLeuAspSerArgProPheValGluTrpAsnThrSer 40
DB 244 AGTGAAGCGGAAAAAGTGTCTGCTAATTGATGATGACCGGCACTTTTGTAATACATCATCC 303
QY 41 HisIleLeuGluAlaIleAsnIleAsnCysSerLysLeuMetLysArgArgLeuGlnGln 60
DB 304 CACATTTTGGAGCCATTAATATCAACTGCTCCAGCTTATGAAAGCGAAGGTGCAACAG 363
QY 61 AspLysValLeuIleThrGluLeuLysGlnHisSerAlaLysHisLysValAspIleAsp 80
DB 364 GACAAAGTATTATTAACAGAGCTCATCAAGCAATTCAGGAAACATTAAGCTTGACATTGAT 423
QY 81 CysSerGlnLysValValValTrpAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB 424 TGCAGTCAGAAAGGTGTGATTGATACATCAAGAGCTCCCAAGATGTTGCTCTCTCTTCA 483
QY 101 AspCysPheLeuThrValLeuLeuGlyLysLeuGlyLysSerPheAsnSerValHisLeu 120
DB 484 GACTCTTTTCTCAGTACTTCTGGTAACTGAGAAAGAGCTTCAACTCTGTTCACCTG 543
QY 121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLysGlyLysSer 140
DB 544 CTTCGAGGTGGGTTTGTGAGTTCTCTGCTGTTTCCCTGGCCCTTGTAAGGAAAAATCC 603
QY 141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB 604 ACTTAGTCCCTTACCTGATTTCTCAGCTTGTCTTACTGTGTCACATTTGGGCCAAC 663
QY 161 ArgIleLeuProAsnLeuTrpLeuGlyCysGlnArgAspValLeuAsnLysGluLeuIle 180
DB 664 GCAATTTCTCCCAATCTTATCTTGGCTGCCAGAGAAATGCTCCACAAAGAGCTGAT- 722
QY 181 -GlnGlnAsnGlyIleGlyTrpValLeuAsnAlaSerTrp-ThrCysProLysProAsp 200
DB 723 GCACAGAAATGAGTATGTTATGTTAATGCAAGAA-TACTCTGCCAAAGCTGACT 781
QY 200 helLeProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIle 220
DB 782 TTATCCCGAGTCTCATTTCTGCTGCTGCTGATGATGACACTTTTGGAGAAATTT 841
QY 220 euProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsnGlyCys 240
DB 842 TGCCTGTGTTGACCAATCTGATGATTTCAATGAAAGCAAAAGCTTCAATGAGATGG 901
QY 240 alleuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTrpIle 260
DB 902 TTCTAGTCACTGTTTAGCTGGGATCTCCGCTCCGACACATCGCTATCGCTACATA 961
QY 260 eLysArgMetAspMetSerLeuAspGluAlaTrpArgPheValLysGluLysArgPro 280
DB 962 TGAAGAGATGACATGCTTTATGATGAAGCTTACAGATTTTGGAAAGAAAAAGACCTTA 1021
QY 280 hrlLeuSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTrpGluLysLysIleVal 300
DB 1022 CTATATCTCAAACTTCAATTTTCTGCGCAACTCTGACATATGAGAAAGATTTAACA 1081
QY 300 snGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlyLysPro 320
DB 1082 ACCAGACTGAGCATCAGGGCCCAAGAGCAAACTCAAGCTGTGCACCTGGAGAACCA 1141
QY 320 snGluProValProAlaValSerGluGlyGlyLysSerGlnThrProLeuSerPro 340
DB 1142 ATGAACTTGTCTGCTGCTGCTGAGAGGGGAGCAAGAAAGCGCCCTCACTCAC 1201
QY 340 roCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSer 360
DB 1202 CTGTGCGGACTGTGTACTCAGAGGACGACGACAAAGCGCTGATCCGCGACAGG 1261
QY 360 alProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAla 380
DB 1262 TGCCAGCGGTGCCAGCGTGCAGCGCTGCTGTGAGAGGACGCGCGTGTAGAGGCGC 1321

```

QY 380 euserGlyLeuHsLeuSerAlaAspArgLeuGluAspSerAenLyLeuLysArgSerP 400
    |||
Db 1322 TCAGTGGCTGCACTGCTCGCAGACAGGCTGGAAAGCAATTAAGCTCAAGCTTCC 1381
QY 400 hseSerLeuAspLysSerValSerTyrSerAlaSerMetAlaAspLeuHsGlyP 420
    |||
Db 1382 TCTCTCTGGATTCATATCATGTTTCAATTCACCCAGCATGGAGCATCTTACATGGCT 1441
QY 420 hseSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAspGlyThA 440
    |||
Db 1442 TCTCCCTCATCAGAAAGATGCTTTGGAAATACCAAACTTCCACTGATCTGGATGGAGCA 1501
QY 440 enLyLeuGlyCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlnThrSerP 460
    |||
Db 1502 ACAAGCTAAGCCAGCTTCTCCCTGTTTCAGAGAACTATCGAGAGACTCCCGAAACCAAGTC 1561
QY 460 roAspLysGluGluAlaSerLysProLysLeuGlnThrAlaArgProSerAspSerG 480
    |||
Db 1562 CTGATTAAGAGAGAGAGCCAGATCCCAAGAGCTGCAGACCCGAGCTTCAGACAGCC 1621
QY 480 InSerLysArgLeuHsSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
    |||
Db 1622 AGAGCAAGCGATTCATTCGCTGAGAACAGCAGCAGTGGCAGCCGCGCAGAGGCTCCCTTT 1681
QY 500 euserProLeuHsArgSerGlySerValGluAspAspThrHsThrSerPheLeuPheG 520
    |||
Db 1682 TATCTCCTGATCCAGAGGAGGAGGAGGAGGAGCAATTAACACACAGCTTCTTTTGG 1741
QY 520 LysSerThrSerGlnGlnHsLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTyrP 540
    |||
Db 1742 GCTTTCACACAGCCAGCAGCAGCTTCAAGAGTCTGCTGGCTGGGCTTMAAGGCTGGC 1801
QY 540 isSerAspLysLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrTyrPheA 560
    |||
Db 1802 ACTCGGATATCTTGGCCCCCAGACCTTACCTTCCCTGACACAGCAGCTGATTTTG 1861
QY 560 lathrGluSerSerHsPheTyrSerAlaSerAlaLeuTyrGlyGlyLysSerAlaSerTyrS 580
    |||
Db 1862 CCACAGAGTCTTCACTTCTTACTGCTGCTCAGCCATCTAAGAGGAGGAGGAGGAGTACT 1921
QY 580 eSerLysSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArgA 600
    |||
Db 1922 CTGCCCTACAGCTGACAGCTGCCCACTTGGCGGAGACCAAGCTATTTCTGTGCGAGGC 1981
QY 600 rGlnLysProSerAspArgAlaAspSerArgArgSerTyrPheGlnGluSerProPheG 620
    |||
Db 1982 GGCAGAGCCAGAGTGCAGAGCTGACTCGCGCGGAGCTGCATGAAGAGAGGCCCTTTG 2041
QY 620 LysLeuGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerLysMetSerGluA 640
    |||
Db 2042 AAAAGAGATTAAACCCAGAGAGCTGCCAAATGGAATTTGAGAGACATCATGTTCAGAGA 2101
QY 640 enArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMetG 660
    |||
Db 2102 ACAGGTACAGGAGAGAGCTGGGAAAGTGGGCACTGACTTACTTTTGGGAGAGATGG 2161
QY 660 LysLeuLeuGluValSer 665
    |||
Db 2162 AATTCATTGAGGTCTCC 2178
    |||

```

```

; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; FILE REFERENCE: 21402-258
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 255
; LENGTH: 2200
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-012-255
Alignment Scores:
Pred. No.: 0
Score: 408.00
Percent Similarity: 99.32%
Beet Local Similarity: 99.32%
Query Match: 61.35%
DB: 13
Length: 2200
Matches: 435
Conservative: 0
Mismatch: 0
Indels: 3
Gaps: 0
US-10-029-345a-109 (1-665) x US-10-072-012-255 (1-2200)
QY 231 GlnLysAlaLysAlaSerAsnGlyCysValLeuValHsCysLeuAlaGlyLysSerArg 250
    |||
Db 793 GAGAAAGCAAAAGCCCTCCATGATGATGTCTTAGTGCACCTGTTTACTGGATCTCCGC 852
QY 251 SerAlaThrLysAlaLysAlaLysAlaTyrLysMetLysArgMetAspMetSerLeuAspGluA 270
    |||
Db 853 TCCGCACCATCGTATCGCTACATCATGAGAGAGATGAGATGCTTAAAGTGAAGCT 912
QY 271 Tyr---ArgPheValLysGluLysArgProThrLysSerProAsnPheAsnPheLeuGly 289
    |||
Db 913 TACAGAGATTGTGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 972
QY 290 GlnLeuLeuAspTyrGluLysLysLysLysLysLysLysLysLysLysLysLysLysLys 309
    |||

```



```
; APPLICANT: LINCOLN, Stephen E.
; APPLICANT: JACKSON, Stuart
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: FT-1113 PCT
; CURRENT APPLICATION NUMBER: US/10/220,120
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,774;
; 60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
; 60/184,769; 60/184,766; 60/184,837; 60/184,697; 60/184,841;
; 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
; 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
; 60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
; 60/205,324; 60/205,286
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-05-17; 2000-05-12; 2000-05-12; 2000-05-12; 2000-05-15;
; 2000-05-16; 2000-05-16; 2000-05-16; 2000-05-16; 2000-05-17;
; 2000-05-17; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PERL Program
; SEQ ID NO 28
; LENGTH: 2476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20040048253A1 LI:1085273.2:2000MAY01
US-10-220-120-28

Alignment Scores:
Pred. No.: 1,11e-09 Length: 2476
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3,01% Indels: 0
DB: 13 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-220-120-28 (1-2476)

Qy 242 ValHhCyeLnuAaglylleserAgsrAlaThrlleAlleAlleAllyllemetlys 261
Db 933 GTCCTGCTGCTGGCTGGCATCTCCGCTCTGCCACCAATCCGCTCATCATGAGG 992

RESULT 19
US-09-925-299-69
; Sequence 69, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 1848
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: n equals a,t,g, or c
; APPLICANT: Bret P. Monia
; APPLICANT: Lex M. Cowart
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 6 EXPRESSION
```

```
; LOCATION: (1847)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-69

Alignment Scores:
Pred. No.: 0.104 Length: 1848
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 9 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-925-299-69 (1-1848)

Qy 240 ValHhCyeLnuAaglylleserAgsr 251
Db 290 GTCCTGTAATGCTGCTGGCATTCACCGCTCA 325

RESULT 20
US-09-925-299-69
; Sequence 69, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 1848
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: n equals a,t,g, or c
; APPLICANT: Bret P. Monia
; APPLICANT: Lex M. Cowart
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 6 EXPRESSION
```

```
/ FILE REFERENCE: PTS-0009
/ CURRENT APPLICATION NUMBER: US/10/199,221
/ CURRENT FILING DATE: 2002-07-18
/ NUMBER OF SEQ ID NOS: 101
/ SEQ ID NO 11
/ LENGTH: 1952
/ TYPE: DNA
/ ORGANISM: H. sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (352)...(1059)
US-10-199-221-11

Alignment Scores:
Pred. No.: 0,109      Length: 1952
Score: 12.00      Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.80%      Indels: 0
DB: 16      Gaps: 0

US-10-029-345A-109 (1-665) x US-10-199-221-11 (1-1952)

QY 240 ValLeuValHisCysLeuAlaGlyIleSerArgSer 251
Db 778 GTCCTGGTACATGCTTGCTGCGCATTAAGCCGCTCA 813

RESULT 22
US-10-072-012-265
/ Sequence 265, Application US/10072012
/ Publication No. US20040033493A1
/ GENERAL INFORMATION:
/ APPLICANT: Tcherenev, Velizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zerhusen, Bryan
/ APPLICANT: Patturajan, Meena
/ APPLICANT: Shimkets, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangolli, Esha
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimир Y.
/ APPLICANT: Coleman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grose, William M.
/ APPLICANT: Alsobrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072,012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 60/266,767
```

```
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: 60/267,057
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/266,975
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/267,459
/ PRIOR FILING DATE: 2001-02-08
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1391
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 265
/ LENGTH: 2039
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-072-012-265

Alignment Scores:
Pred. No.: 0,114      Length: 2039
Score: 12.00      Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.80%      Indels: 0
DB: 13      Gaps: 0

US-10-029-345A-109 (1-665) x US-10-072-012-265 (1-2039)

QY 281 lIeSerProAsnPhaeAsnPhleuGlyGlnIleuLeu 292
Db 952 ATCTGCCCACTCACTTCTCTGCGCAGCTGCTG 987

RESULT 23
US-10-152-319A-2041
/ Sequence 2041, Application US/10152319A
/ Publication No. US20040072160A1
/ GENERAL INFORMATION:
/ APPLICANT: Mendrick, Donna
/ APPLICANT: Porter, Mark
/ APPLICANT: Johnson, Kory
/ APPLICANT: Higgs, Brandon
/ APPLICANT: Castle, Arthur
/ APPLICANT: Elashoff, Michael
/ TITLE OF INVENTION: Molecular Toxicology Modeling
/ FILE REFERENCE: 44921-5089-US
/ CURRENT APPLICATION NUMBER: US/10/152,319A
/ CURRENT FILING DATE: 2002-05-22
/ PRIOR APPLICATION NUMBER: US 60/292,335
/ PRIOR FILING DATE: 2001-05-22
/ PRIOR APPLICATION NUMBER: US 60/297,523
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/298,925
/ PRIOR FILING DATE: 2001-06-19
/ PRIOR APPLICATION NUMBER: US 60/303,810
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/303,807
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/303,808
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/315,047
/ PRIOR FILING DATE: 2001-08-28
/ PRIOR APPLICATION NUMBER: US 60/324,928
/ PRIOR FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: US 60/330,867
/ PRIOR FILING DATE: 2001-11-01
/ PRIOR APPLICATION NUMBER: US 60/330,462
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 2221
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2041
/ LENGTH: 2104
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
/ FEATURE:
```

```
; OTHER INFORMATION: Genbank Accession No. NM_053883
US-10-152-319A-2041

Alignment Scores:
Pred. No.: 0.117      Length: 2104
Score: 12.00          Matches: 12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.80%      Indels: 0
DB: 12                Gaps: 0

US-10-029-345A-109 (1-665) x US-10-152-319A-2041 (1-2104)
QY 240 ValLeuValHicCySeLeuAaGlyIleSeArGser 251
Db 1225 GTCCTGTCATTCCTTGCGGCGCATCAGCCGCTCC 1260

RESULT 24
US-10-388-934-654
; Sequence 654, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 654
; LENGTH: 2104
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547A1way rat)
US-10-388-934-654

Alignment Scores:
Pred. No.: 0.117      Length: 2104
Score: 12.00          Matches: 12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.80%      Indels: 0
DB: 16                Gaps: 0

US-10-029-345A-109 (1-665) x US-10-388-934-654 (1-2104)
QY 240 ValLeuValHicCySeLeuAaGlyIleSeArGser 251
Db 1225 GTCCTGTCATTCCTTGCGGCGCATCAGCCGCTCC 1260

RESULT 25
US-10-191-803-270
; Sequence 270, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: BLASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
```

```
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 270
; LENGTH: 2104
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_053883
US-10-191-803-270

Alignment Scores:
Pred. No.: 0.117      Length: 2104
Score: 12.00          Matches: 12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.80%      Indels: 0
DB: 16                Gaps: 0

US-10-029-345A-109 (1-665) x US-10-191-803-270 (1-2104)
QY 240 ValLeuValHicCySeLeuAaGlyIleSeArGser 251
Db 1225 GTCCTGTCATTCCTTGCGGCGCATCAGCCGCTCC 1260

RESULT 26
US-10-305-720-1135
; Sequence 1135, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expressi
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1135
; LENGTH: 2109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Genbank ID No. US20040010136A1 g1418933
US-10-305-720-1135

Alignment Scores:
Pred. No.: 0.117      Length: 2109
Score: 12.00          Matches: 12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.80%      Indels: 0
DB: 16                Gaps: 0

US-10-029-345A-109 (1-665) x US-10-305-720-1135 (1-2109)
QY 240 ValLeuValHicCySeLeuAaGlyIleSeArGser 251
Db 1216 GTCCTGTCATTCCTTGCGGCGCATCAGCCGCTCA 1251

RESULT 27
US-10-641-643-946
; Sequence 946, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
```

```
Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
GENE EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/641,643
FILING DATE: 14-Aug-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 946:
SEQUENCE CHARACTERISTICS:
LENGTH: 2109 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1418933
SEQUENCE DESCRIPTION: SEQ ID NO: 946 :
US-10-641-643-946

Alignment Scores:
Pred. No.: 0.117 Length: 2109
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 17 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-641-643-946 (1-2109)
QY 240 ValIeuValHISCySleuAlaGlyIleSerArgSer 251
Db 1216 GTCTTGTAATTGCTTGCTGGCATTTAGCCGCTCA 1251

RESULT 28
US-10-342-887-698
; Sequence 698, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
```

```
PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 698
; LENGTH: 2390
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-698

Alignment Scores:
Pred. No.: 0.131 Length: 2390
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 13 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-172-118-698 (1-2390)
QY 240 ValIeuValHISCySleuAlaGlyIleSerArgSer 251
Db 1216 GTCTTGTAATTGCTTGCTGGCATTTAGCCGCTCA 1251

RESULT 30
US-10-175-523-183
; Sequence 183, Application US/10175523

Alignment Scores:
Pred. No.: 0.131 Length: 2390
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 13 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-172-118-698 (1-2390)
QY 240 ValIeuValHISCySleuAlaGlyIleSerArgSer 251
Db 1216 GTCTTGTAATTGCTTGCTGGCATTTAGCCGCTCA 1251

RESULT 30
US-10-175-523-183
; Sequence 183, Application US/10175523
```

Publication No. US20030096264A1
GENERAL INFORMATION:
APPLICANT: Brockman, Jeffrey
APPLICANT: Evans, David
APPLICANT: Hook, Derek
APPLICANT: Klimczak, Leszek
APPLICANT: Laeng, Pascal
APPLICANT: Paley, Michael
APPLICANT: Rajan, Prithi
TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPTS)
FILE REFERENCE: 3235/1795-US3
CURRENT APPLICATION NUMBER: US/10/175,523
CURRENT FILING DATE: 2002-06-18
PRIOR APPLICATION NUMBER: US 60/299,151
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: US 60/317,828
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US 60/325,150
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/333,047
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: US 60/349,936
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US 60/361,834
PRIOR FILING DATE: 2002-03-04
NUMBER OF SEQ ID NOS: 197
SOFTWARE: PatentIn version 3.1
SEQ ID NO 183
LENGTH: 2390
TYPE: DNA
ORGANISM: Homo sapiens
US-10-175-523-183

Alignment Scores:
Pred. No.: 0.131 Length: 2390
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 15 Gaps: 0

US-10-029-345a-109 (1-665) x US-10-175-523-183 (1-2390)
QY 240 ValLeuValHicCyLeuAlaGlyTlleSerArgSer 251
DB 1216 GTCCTGTACATTCCTGCTGCGCATTCAGCCGCTCA 1251

RESULT 31
US-10-199-221-4
Sequence 4, Application US/10199221
Publication No. US20040014048A1
GENERAL INFORMATION:
APPLICANT: Brett P. Monia
APPLICANT: Lex M. Cowart
APPLICANT: Kenneth W. Dobbie
TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 6 EXPRESSION
FILE REFERENCE: FTS-0009
CURRENT APPLICATION NUMBER: US/10/199,221
CURRENT FILING DATE: 2002-07-18
NUMBER OF SEQ ID NOS: 101
SEQ ID NO 4
LENGTH: 2390
TYPE: DNA
ORGANISM: H. sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (352)...(1497)
US-10-199-221-4

Alignment Scores:
Pred. No.: 0.131 Length: 2390
Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 16 Gaps: 0

US-10-029-345a-109 (1-665) x US-10-199-221-4 (1-2390)

QY 240 ValLeuValHicCyLeuAlaGlyTlleSerArgSer 251
DB 1216 GTCCTGTACATTCCTGCTGCGCATTCAGCCGCTCA 1251

RESULT 32
US-10-717-597-4
Sequence 4, Application US/10717597
Publication No. US20040110221A1
GENERAL INFORMATION:
APPLICANT: Wyeth
APPLICANT: Burczynski, Michael E.
APPLICANT: Twine, Natalie C.
APPLICANT: Dornier, Andrew J.
APPLICANT: Trepicchio, William L.
APPLICANT: Slonim, Donna K.
APPLICANT: Stover, Jennifer A.
TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
FILE REFERENCE: AM101080L
CURRENT APPLICATION NUMBER: US/10/717,597
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: US 60/459,782
PRIOR FILING DATE: 2003-04-03
PRIOR APPLICATION NUMBER: US 60/427,982
PRIOR FILING DATE: 2002-11-21
NUMBER OF SEQ ID NOS: 4904
SOFTWARE: PatentIn version 3.2
SEQ ID NO 4
LENGTH: 2390
TYPE: DNA
ORGANISM: Homo sapiens
US-10-717-597-4

Alignment Scores:
Pred. No.: 0.131 Length: 2390
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 17 Gaps: 0

US-10-029-345a-109 (1-665) x US-10-717-597-4 (1-2390)

QY 240 ValLeuValHicCyLeuAlaGlyTlleSerArgSer 251
DB 1216 GTCCTGTACATTCCTGCTGCGCATTCAGCCGCTCA 1251

RESULT 33
US-10-133-937-48
Sequence 48, Application US/10133937
Publication No. US20030207278A1
GENERAL INFORMATION:
APPLICANT: Khan, Javed
APPLICANT: Ringner, Markus
APPLICANT: Peterson, Carsten
APPLICANT: Meltzer, Paul
TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING, DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
FILE REFERENCE: 11613.56US01
CURRENT APPLICATION NUMBER: US/10/133,937
CURRENT FILING DATE: 2002-11-04
NUMBER OF SEQ ID NOS: 99
SOFTWARE: PatentIn version 3.1
SEQ ID NO 48
LENGTH: 2649
TYPE: DNA

```
/ ORGANISM: Homo sapiens
US-10-133-937-48

Alignment Scores:
Pred. No.: 0.144 Length: 2649
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 16 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-133-937-48 (1-2649)
QY 240 ValieuValHsCysLeuA1agLYIleSerArgSer 251
Db 1153 GTCTTGATCATTTGCTTGCGCATTTAGCCGCTCA 1188

RESULT 34
US-10-159-563-48
/ Sequence 48, Application US/10159563
/ Publication No. US20040009154A1
/ GENERAL INFORMATION:
/ APPLICANT: Khan, Javed
/ APPLICANT: Ringner, Markus
/ APPLICANT: Peterson, Carsten
/ APPLICANT: Meltzer, Paul
/ TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
/ FILE REFERENCE: 11613.56US11
/ CURRENT APPLICATION NUMBER: US/10/159,563
/ PRIOR FILING DATE: 2002-12-09
/ PRIOR APPLICATION NUMBER: US 10/133,937
/ NUMBER OF SEQ ID NOS: 444
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 48
/ LENGTH: 2649
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-159-563-48

Alignment Scores:
Pred. No.: 0.144 Length: 2649
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 16 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-159-563-48 (1-2649)
QY 240 ValieuValHsCysLeuA1agLYIleSerArgSer 251
Db 1153 GTCTTGATCATTTGCTTGCGCATTTAGCCGCTCA 1188

RESULT 35
US-10-184-832-1
/ Sequence 1, Application US/10184832
/ Publication No. US20030022857A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu et al.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
/ TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
/ FILE REFERENCE: MP12001-056P1NM
/ CURRENT APPLICATION NUMBER: US/10/184,832
/ PRIOR FILING DATE: 2002-06-28
/ PRIOR APPLICATION NUMBER: 60/303,250
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 2650
/ TYPE: DNA
```

```
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (280) ... (1422)
US-10-184-832-1

Alignment Scores:
Pred. No.: 0.144 Length: 2650
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 15 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-184-832-1 (1-2650)
QY 240 ValieuValHsCysLeuA1agLYIleSerArgSer 251
Db 1144 GTCTTGATCATTTGCTTGCGCATTTAGCCGCTCT 1179

RESULT 36
US-09-971-392-69
/ Sequence 69, Application US/09971392
/ Publication No. US20030134283A1
/ GENERAL INFORMATION:
/ APPLICANT: Peterson, David P.
/ APPLICANT: Pearson, Cecelia I.
/ APPLICANT: Cocks, Benjamin G.
/ TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
/ FILE REFERENCE: PA-0029 US
/ CURRENT APPLICATION NUMBER: US/09/971,392
/ PRIOR FILING DATE: 2001-10-03
/ PRIOR APPLICATION NUMBER: 60/237,652
/ NUMBER OF SEQ ID NOS: 260
/ SOFTWARE: PERL Program
/ SEQ ID NO 69
/ LENGTH: 2792
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Template ID: 149914.21
/ LOCATION: 2789
/ OTHER INFORMATION: a, t, c, g, or other
US-09-971-392-69

Alignment Scores:
Pred. No.: 0.151 Length: 2792
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 10 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-971-392-69 (1-2792)
QY 240 ValieuValHsCysLeuA1agLYIleSerArgSer 251
Db 1314 GTCTTGATCATTTGCTTGCGCATTTAGCCGCTCA 1349

RESULT 37
US-10-044-090-330
/ Sequence 330, Application US/10044090
/ Publication No. US20020137081A1
/ GENERAL INFORMATION:
/ APPLICANT: Olga Bandman
/ TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
/ FILE REFERENCE: PA-0028 US
/ CURRENT APPLICATION NUMBER: US/10/044,090
/ PRIOR FILING DATE: 2002-01-09
/ NUMBER OF SEQ ID NOS: 850
```

```

: SOFTWARE: PERL Program
: SEQ ID NO 330
: LENGTH: 2794
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID NO. US20020137081A1 060151CB1
US-10-044-090-330

Alignment Scores:
Pred. No.: 0.151 Length: 2794
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 14 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-044-090-330 (1-2794)
OY 240 VALLEUVAlHIsCYsLEuAlaGlytLeserAArgSer 251
|||||
Db 1314 GTCTTGTAACATTGCTGCTGGCTGGCAATTAAGCGCTCA 1349
|||||

RESULT 38
US-10-199-221-12
: Sequence 12, Application US/10199221
: Publication No. US20040014048A1
: GENERAL INFORMATION:
: APPLICANT: Brett P. Monia
: APPLICANT: Lex M. Cowbert
: APPLICANT: Kenneth W. Dobbie
: TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 6 EXPRESSION
: FILE REFERENCE: Pts-0009
: CURRENT APPLICATION NUMBER: US/10/199,221
: CURRENT FILING DATE: 2002-07-18
: NUMBER OF SEQ ID NOS: 101
: SEQ ID NO 12
: LENGTH: 7001
: TYPE: DNA
: ORGANISM: H. sapiens
: FEATURE:
US-10-199-221-12

Alignment Scores:
Pred. No.: 0.35 Length: 7001
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 16 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-199-221-12 (1-7001)
OY 240 VALLEUVAlHIsCYsLEuAlaGlytLeserAArgSer 251
|||||
Db 4496 GTCTTGTAACATTGCTGCTGGCTGGCAATTAAGCGCTCA 4531
|||||

RESULT 39
US-10-085-117-241/c
: Sequence 241, Application US/10085117
: Publication No. US2003023234A1
: GENERAL INFORMATION:
: APPLICANT: Moritz, David W.
: APPLICANT: Engelhard, Eric K.
: TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
: FILE REFERENCE: 529452000121
: CURRENT APPLICATION NUMBER: US/10/085,117
: CURRENT FILING DATE: 2002-02-27
: PRIOR APPLICATION NUMBER: US 09/798,586
: PRIOR FILING DATE: 2001-03-02
: NUMBER OF SEQ ID NOS: 361
: SOFTWARE: FASTSEQ for Windows Version 4.0

```

```

; SEQ ID NO 241
; LENGTH: 165961
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1) ... (165961)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-241

Alignment Scores:
Pred. No.: 664 Length: 165961
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.50% Indels: 0
DB: 16 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-085-117-241 (1-165961)
Oy 300 AsnGlnThrGlyAlaSerGlyProLysSer 309
|||||
Db 136778 AACCAACAGAGGACGACTGGCCCAAGAGC 136749

RESULT 40
US-10-006-285-423
; Sequence 423, Application US/10006285
; Publication No. US20030165854A1
; GENERAL INFORMATION:
; APPLICANT: Mary Jane Cunningham
; APPLICANT: Matthew R. Kaber
; TITLE OF INVENTION: MARKER GENES RESPONDING TO TREATMENT WITH TOXINS
; FILE REFERENCE: PA-0039 US
; CURRENT APPLICATION NUMBER: US/10/006,285
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 514
; SOFTWARE: PERL Program
; SEQ ID NO 423
; LENGTH: 1571
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030165854A1 332413.6
US-10-006-285-423

Alignment Scores:
Pred. No.: 96.5 Length: 1571
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 15 Gaps: 0

US-10-029-345A-109 (1-665) x US-10-006-285-423 (1-1571)
Oy 495 AlaGlnGSerLeuLeuSerProLeu 503
|||||
Db 104 GCCCAGCGTTCCTCTCTCTCCACTT 130

RESULT 41
US-09-736-457-801
; Sequence 801, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane

```

```

; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736.457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-801

Alignment Scores:
Pred. No.: 99.2          Length: 1619
Score: 9.00             Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35%      Indels: 0
DB: 9                   Gaps: 0

US-10-029-345A-109 (1-665) x US-09-736-457-801 (1-1619)

QY 246 AlaGlyIleSerArgSerAlaThrIle 254
DB 569 GCGGGCATCTCGGGGTGGCCACCATC 595

RESULT 42
US-09-902-941-801
; Sequence 801, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnierakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-801

Alignment Scores:
Pred. No.: 99.2          Length: 1619
Score: 9.00             Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35%      Indels: 0
DB: 9                   Gaps: 0

US-10-029-345A-109 (1-665) x US-09-902-941-801 (1-1619)

QY 246 AlaGlyIleSerArgSerAlaThrIle 254
DB 569 GCGGGCATCTCGGGGTGGCCACCATC 595

RESULT 43
US-10-029-345A-109 (1-665) x US-09-902-941-801 (1-1619)

QY 246 AlaGlyIleSerArgSerAlaThrIle 254
DB 569 GCGGGCATCTCGGGGTGGCCACCATC 595

RESULT 43
```

```

US-09-849-626-801
; Sequence 801, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tonglong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-626-801

Alignment Scores:
Pred. No.: 99.2          Length: 1619
Score: 9.00             Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35%      Indels: 0
DB: 9                   Gaps: 0

US-10-029-345A-109 (1-665) x US-09-849-626-801 (1-1619)

QY 246 AlaGlyIleSerArgSerAlaThrIle 254
DB 569 GCGGGCATCTCGGGGTGGCCACCATC 595

RESULT 44
US-10-283-017-801
; Sequence 801, Application US/10283017
; Publication No. US20030211510A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C20
; CURRENT APPLICATION NUMBER: US/10/283,017
; CURRENT FILING DATE: 2002-10-28
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-283-017-801

Alignment Scores:
Pred. No.: 99.2          Length: 1619
Score: 9.00             Matches: 9
Percent Similarity: 100.00% Conservative: 0
```


Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 13 Gaps: 0
US-10-029-345A-109 (1-665) x US-10-283-017-801 (1-1619)
QY 246 AAGGYYIleSerArgSerAlaThrIle 254
DB 569 GCGGGCATCTCGCGGTCCGCGCACACATC 595
RESULT 45
US-10-017-754-801
Sequence 801, Application US/10017754
GENERAL INFORMATION:
APPLICANT: Henderson, Robert A.
APPLICANT: Wang, Tonglong
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Johnson, Jeffrey C.
APPLICANT: Retter, Marc W.
APPLICANT: Marnerakis, Margarita
APPLICANT: Carter, Derrick
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: McNabb, Andria
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C18
CURRENT APPLICATION NUMBER: US/10/017,754
CURRENT FILING DATE: 2001-10-29
NUMBER OF SEQ ID NOS: 2004
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 801
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapiens
US-10-017-754-801
Alignment Scores:
Pred. No.: 99.2 Length: 1619
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 15 Gaps: 0
US-10-029-345A-109 (1-665) x US-10-017-754-801 (1-1619)
QY 246 AAGGYYIleSerArgSerAlaThrIle 254
DB 569 GCGGGCATCTCGCGGTCCGCGCACACATC 595
RESULT 46
US-10-113-872-801
Sequence 801, Application US/10113872
Publication No. US20030170255A1
GENERAL INFORMATION:
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Sleath, Paul R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Derrick
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C19
CURRENT APPLICATION NUMBER: US/10/113,872
CURRENT FILING DATE: 2002-03-28
NUMBER OF SEQ ID NOS: 2011
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 801

LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapiens
US-10-113-872-801
Alignment Scores:
Pred. No.: 99.2 Length: 1619
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 15 Gaps: 0
US-10-029-345A-109 (1-665) x US-10-113-872-801 (1-1619)
QY 246 AAGGYYIleSerArgSerAlaThrIle 254
DB 569 GCGGGCATCTCGCGGTCCGCGCACACATC 595
RESULT 47
US-09-971-392-5
Sequence 5, Application US/09971392
Publication No. US20030134283A1
GENERAL INFORMATION:
APPLICANT: Peterson, David P.
APPLICANT: Pearson, Cecelia I.
APPLICANT: Cocks, Benjamin G.
TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
FILE REFERENCE: PA-0029 US
CURRENT APPLICATION NUMBER: US/09/971,392
CURRENT FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: 60/237,652
PRIOR FILING DATE: 2000-10-03
NUMBER OF SEQ ID NOS: 260
SOFTWARE: PERL Program
SEQ ID NO 5
LENGTH: 1632
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Template ID: 990840.12
US-09-971-392-5
Alignment Scores:
Pred. No.: 99.9 Length: 1632
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 10 Gaps: 0
US-10-029-345A-109 (1-665) x US-09-971-392-5 (1-1632)
QY 246 AAGGYYIleSerArgSerAlaThrIle 254
DB 1106 GCGGGCATCTCGCGGTCCGCGCACACATC 1132
RESULT 48
US-10-002-600-51
Sequence 51, Application US/10002600
Publication No. US20020137077A1
GENERAL INFORMATION:
APPLICANT: Hopkins, Christopher M.
APPLICANT: Peterson, David P.
APPLICANT: Cocks, Benjamin G.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: GENES REGULATED IN ACTIVATED T CELLS
FILE REFERENCE: PA-0042 US
CURRENT APPLICATION NUMBER: US/10/002,600
CURRENT FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: 60/243,521
PRIOR FILING DATE: 2000-10-25

```
/ NUMBER OF SEQ ID NOS: 116
/ SOFTWARE: PERL Program
/ SEQ ID NO 51
/ LENGTH: 1685
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Template ID: 235631.2
US-10-002-600-51

Alignment Scores:
Pred. No.: 103          Length: 1685
Score: 9.00           Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35%      Indels: 0
DB: 14                Gaps: 0

US-10-029-345A-109 (1-665) x US-10-002-600-51 (1-1685)

QY 246 AlaglylleserArgSerAlaThrIle 254
DB 862 GCGGGTATCTCGCGCTCTGCCACCATC 886

RESULT 49
US-10-342-887-1034
/ Sequence 1034, Application US/10342887
/ Publication No. US20040058340A1
/ GENERAL INFORMATION:
/ APPLICANT: Dai, Hongyue
/ APPLICANT: He, Yudong
/ APPLICANT: Linsley, Peter S.
/ APPLICANT: Mao, Mao
/ APPLICANT: Roberts, Christopher J.
/ APPLICANT: Van 't Veer, Laura Johanna
/ APPLICANT: Van de Vijver, Marc J.
/ TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
/ FILE REFERENCE: 9301-188-999
/ CURRENT APPLICATION NUMBER: US/10/342,887
/ CURRENT FILING DATE: 2003-01-15
/ PRIOR APPLICATION NUMBER: 60/298,918
/ PRIOR FILING DATE: 2001-06-18
/ PRIOR APPLICATION NUMBER: 60/380,710
/ PRIOR FILING DATE: 2002-05-14
/ PRIOR APPLICATION NUMBER: 10/172,118
/ PRIOR FILING DATE: 2002-06-14
/ NUMBER OF SEQ ID NOS: 2699
/ SEQ ID NO 1034
/ LENGTH: 1702
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-342-887-1034

Alignment Scores:
Pred. No.: 104          Length: 1702
Score: 9.00           Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35%      Indels: 0
DB: 13                Gaps: 0

US-10-029-345A-109 (1-665) x US-10-342-887-1034 (1-1702)

QY 246 AlaglylleserArgSerAlaThrIle 254
DB 860 GCGGGTATCTCGCGCTCTGCCACCATC 886

RESULT 50
US-10-172-118-1034
/ Sequence 1034, Application US/10172118
/ Publication No. US20030224374A1
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Dai, Hongyue
/ APPLICANT: He, Yudong
/ APPLICANT: Linsley, Peter
/ APPLICANT: Mao, Mao
/ APPLICANT: Roberts, Chris
/ APPLICANT: Van 't Veer, Laura
/ APPLICANT: Van de Vijver, Marc
/ APPLICANT: Bernards, Rene
/ TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
/ FILE REFERENCE: 9301-175-999
/ CURRENT APPLICATION NUMBER: US/10/172,118
/ CURRENT FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: 60/380,770
/ PRIOR FILING DATE: 2002-05-14
/ NUMBER OF SEQ ID NOS: 2699
/ SEQ ID NO 1034
/ LENGTH: 1702
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: NM_004418
/ DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1034

Alignment Scores:
Pred. No.: 104          Length: 1702
Score: 9.00           Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35%      Indels: 0
DB: 13                Gaps: 0

US-10-029-345A-109 (1-665) x US-10-172-118-1034 (1-1702)

QY 246 AlaglylleserArgSerAlaThrIle 254
DB 860 GCGGGTATCTCGCGCTCTGCCACCATC 886
```

Search completed: June 22, 2004, 10:25:07
Job time : 835 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2004, 06:30:22 ; Search time 122 Seconds

(without alignments)
3024.937 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665

Sequence: 1 MAHEMIGTIVTERLVALLE.....LKVSGSGSFGSGSMETIEVS 665

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 682709 seqs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 92308

Minimum DB seq length: 1419

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US10029345.rn1 21062004.164401.29776/APP QUERY.fastc_1.839
-DB=Issued Patents NA -QFMT=fastc -SUFFIX=oligo.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=humand0.cdi
-LIST=75 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=1419 -MAXLEN=2000000000
-USER=US10029345.QCEN_1_1.56.rn1 21062004.164401.29776 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
6: /cgn2_6/prodata/2/ina/5B_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	71.0	1998	4	US-09-816-494-3
2	472	71.0	3544	4	US-09-816-494-1
3	20	3.0	2377	4	US-09-920-668-3
4	12	1.8	2109	4	US-09-016-434-1135
5	12	1.8	2109	4	US-09-016-434-1135
6	9	1.4	1619	4	US-09-023-655-946
7	9	1.4	1619	4	US-09-702-705-801
8	9	1.4	1619	4	US-09-736-457-801
9	9	1.4	1619	4	US-09-614-124B-801
10	9	1.4	1619	4	US-09-589-184-825
11	9	1.4	1619	4	US-09-589-184-801
12	9	1.4	1993	2	US-08-990-379-1

13	9	1.4	2000	4	US-09-016-434-1291	Sequence 1291, App
14	9	1.4	2064	4	US-09-702-705-825	Sequence 825, App
15	9	1.4	2064	4	US-09-736-457-825	Sequence 825, App
16	9	1.4	2064	4	US-09-614-124B-825	Sequence 825, App
17	9	1.4	2064	4	US-09-671-325-825	Sequence 825, App
18	9	1.4	2064	4	US-09-589-184-825	Sequence 825, App
19	9	1.4	2109	4	US-09-702-705-826	Sequence 826, App
20	9	1.4	2109	4	US-09-736-457-826	Sequence 826, App
21	9	1.4	2109	4	US-09-614-124B-826	Sequence 826, App
22	9	1.4	2109	4	US-09-671-325-826	Sequence 826, App
23	9	1.4	2109	4	US-09-589-184-826	Sequence 826, App
24	9	1.4	2240	4	US-09-016-434-1100	Sequence 1100, App
25	9	1.4	4637	4	US-09-702-705-804	Sequence 804, App
26	9	1.4	4637	4	US-09-736-457-804	Sequence 804, App
27	9	1.4	4637	4	US-09-614-124B-804	Sequence 804, App
28	9	1.4	4637	4	US-09-671-325-804	Sequence 804, App
29	9	1.4	4637	4	US-09-589-184-804	Sequence 804, App
30	9	1.2	1421	5	PCT-US92-08090-3	Sequence 804, App
31	8	1.2	1426	2	US-08-284-465-2	Sequence 3, App1
32	8	1.2	1426	2	US-08-284-465-7	Sequence 2, App1
33	8	1.2	1426	2	PCT-US92-08090-2	Sequence 2, App1
34	8	1.2	1426	2	US-08-284-465-1	Sequence 2, App1
35	8	1.2	1508	3	US-08-236-311-5	Sequence 5, App1
36	8	1.2	1508	3	US-08-457-918-5	Sequence 5, App1
37	8	1.2	1599	2	US-08-417-495-3	Sequence 3, App1
38	8	1.2	1599	2	US-08-284-391B-3	Sequence 3, App1
39	8	1.2	1599	3	US-09-218-950-3	Sequence 3, App1
40	8	1.2	1599	3	PCT-US92-01785-3	Sequence 3, App1
41	8	1.2	1669	2	US-08-916-902A-2	Sequence 3, App1
42	8	1.2	1669	2	US-08-916-902A-2	Sequence 2, App1
43	8	1.2	1728	2	US-08-417-495-1	Sequence 1, App1
44	8	1.2	1728	2	US-08-284-391B-1	Sequence 1, App1
45	8	1.2	1728	2	US-08-284-391B-1	Sequence 1, App1
46	8	1.2	1728	2	US-09-218-950-1	Sequence 1, App1
47	8	1.2	1728	5	PCT-US92-01785-1	Sequence 1, App1
48	8	1.2	1728	5	PCT-US95-00454-1	Sequence 1, App1
49	8	1.2	1742	3	US-08-466-366-3	Sequence 3, App1
50	8	1.2	1742	4	US-09-517-605-7	Sequence 7, App1
51	8	1.2	1742	4	US-08-470-998-1	Sequence 1, App1
52	8	1.2	1742	4	US-09-023-655-1013	Sequence 1013, App
53	8	1.2	1796	3	US-08-328-500-8	Sequence 8, App1
54	8	1.2	1796	3	US-08-477-460B-1	Sequence 1, App1
55	8	1.2	1796	3	US-08-379-516-1	Sequence 1, App1
56	8	1.2	1796	3	US-09-329-916-1	Sequence 1, App1
57	8	1.2	1796	3	US-08-485-372A-1	Sequence 1, App1
58	8	1.2	1796	4	US-09-409-006A-1	Sequence 1, App1
59	8	1.2	1796	5	US-08-484-681-1	Sequence 1, App1
60	8	1.2	1858	3	PCT-US93-07422-1	Sequence 1, App1
61	8	1.2	1858	3	US-08-813-574-1	Sequence 1, App1
62	8	1.2	2303	4	US-09-252-991A-9705	Sequence 9705, App
63	8	1.2	2465	5	US-09-922-146-3	Sequence 3, App1
64	8	1.2	2465	5	PCT-US92-08090-1	Sequence 3, App1
65	8	1.2	2482	3	US-08-477-460B-3	Sequence 3, App1
66	8	1.2	2482	3	US-08-379-516-3	Sequence 3, App1
67	8	1.2	2482	3	US-09-329-916-3	Sequence 3, App1
68	8	1.2	2482	3	US-08-485-372A-3	Sequence 3, App1
69	8	1.2	2482	4	US-09-409-006A-3	Sequence 3, App1
70	8	1.2	2482	4	US-08-484-681-3	Sequence 3, App1
71	8	1.2	2589	5	PCT-US93-07422-3	Sequence 3, App1
72	8	1.2	2589	4	US-08-472-888A-5	Sequence 5, App1
73	8	1.2	2709	1	US-08-021-601-11	Sequence 11, App1
74	8	1.2	2709	1	US-08-082-849B-11	Sequence 11, App1
75	8	1.2	2940	5	PCT-US94-01624-11	Sequence 11, App1
	8	1.2	2940	2	US-08-418-848A-3	Sequence 3, App1

ALIGNMENTS

RESULT 1
US-09-816-494-3
; Sequence 3, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.

; TITLE OF INVENTION: 36692 AND 21117, NOVEL DUAL SPECIFICITY
 ; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: 10448-030002
 ; CURRENT APPLICATION NUMBER: US/09/816,494
 ; CURRENT FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: US 60/191,858
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1998
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-816-494-3

Alignment Scores:

Pred. No.:	0	Length:	1998
Score:	472.00	Matches:	663
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	70.98%	Indels:	4
DB:	4	Gaps:	0

US-10-029-345a-109 (1-665) x US-09-816-494-3 (1-1998)

```

QY      1 MetAlaHisGluMetCllleglYThrGlnlIleValThrGluArgLeuValAlaLeuLeuGlu 20
DB      1 ATGGCCCATGAGATGATGGACATCAAAATGTTACTGACAGGTTGGTGGCTGCTGGAA 60

QY      21 SerGlyThrGluLysValLeuLeuLleAspSerArgProPheValGluTryAenThrSer 40
DB      61 AGTGAACCGGAAAAAGCTGCTAATGATGACCGGCATTTGTGAATATACATATACATCC 120

QY      41 HisIleLeuGluAlaIleAsnIleAsnCySerLysLeuMetLysArgArgLeuGlnGln 60
DB      121 CACATTTTGGAAACCATTAATATCAACTGCTCCAGCTTATGAGGAAAGTTGCAACG 180

QY      61 AspLysValLeuLleThrGluLeuLleGlnHisSerAlaLysHisLysValAspIleAsp 80
DB      181 GACAAAGTGTTAATTAACAGAGCTCATCCAGCATTCAGCGAAACATAGGTTGACATTGAT 240

QY      81 CysSerGluLysValValValValTyrAspGlnSerSerGlnAspValAlaSerLeuSerSer 100
DB      241 TGAAGTCAGAAAGGTGTGATTACGATCAAAAGCTCCCAAGATGTGGCTCTCTCTTCA 300

QY      101 AspCysPheLeuThrValLeuLeuGlyLysLeuGluLysSerPheAsnSerValHisLeu 120
DB      301 GACTGTTTCTCACTGACTTCTGGGTAACCTGGAGAGAGCTTCAACTGTTCACTG 360

QY      121 LeuAlaGlyGlyPheAlaGluPheSerArgCysPheProGlyLeuCySerGluGlyLysSer 140
DB      361 CTTCGAGGTGGGTGGTTCGAGTTCCTCGTGTTCCTCGGCTCGTGGAGAGAAATCC 420

QY      141 ThrLeuValProThrCysIleSerGlnProCysLeuProValAlaAsnIleGlyProThr 160
DB      421 ACTCTAGTCCCTACCTGCACTTCTCAAGCTTCCTTACCTGTTGCCAACAATGGGCAACC 480

QY      161 ArgGlyLeuProAsnLeuLysLeuGlyCysGlnArgAspValLeuAsnLysGlyLeuLle 180
DB      481 CCAATTCCTCCCAATCTTATCTTGCTGCGCCAGCGAGAGTCTCTCAACAAGAGCTGAT- 539

QY      181 -GlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerTyr-ThrCysProLysProAsp 200
DB      540 GCAAGCAAGATGGAGATGGTATGATGTAATGCCAGCAA-TACCTGTCCAAAGCCTGACT 598

QY      200 heIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGluLysIleL 220
DB      599 TTATCCCGAGAGTCTATTTCCGCGCTGTGCTGTGAATGACAGCTTTTGTAGAAAAATTT 658

QY      220 euProTyrLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsnGlyCysV 240
DB      659 TGCCGTGTGTGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCCAATGAGATGTG 718
  
```

```

QY      240 alLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleM 260
DB      719 TTCTAGTCACTGTTTAGTGGGATCTCCCGCTCCGCCACCATTCGCTTACATCA 778

QY      260 eLlyeArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGlyLysArgProT 280
DB      779 TGAAGAGATGAGACATGCTTTAGATGAAAGCTTACAGATTTGCTGAAGAAAAAGACCTA 838

QY      280 hrIleSerProAsnPheAsnPheLeuGlyGlnLeuAsnAspTyrGlyLysIleLys 300
DB      839 CTATATCTCCCAACTTCAATTTTCTGGGCCAACTCTGAACTATGAGAAAGATTAAGA 898

QY      300 snGlnThrGlyAlaSerGlyProLysSerLysLysLeuLeuHisIleuGluLysProA 320
DB      899 ACCAGACTGGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACTGGAGAGACCAA 958

QY      320 snGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeuSerProp 340
DB      959 ATGAACCTGTCTCCCTGCTGTCTTCAGAGGGGTGACAGAAAGGAGACCGCCCTCAGTCCAC 1018

QY      340 roCyAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerV 360
DB      1019 CCTGTGCGGACTCTGCTACTCTCAGAGCGCAGAGACAAAGCCCTGTGATCCCGCCAGCG 1078

QY      360 alProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAla 380
DB      1079 TGCCAGCGGTCCAGCGCTGACGCGCTGCTGTTAGAGACAGCGCCCTGGTACAGCGCG 1138

QY      380 euSerGlyLeuHisIleuSerAlaAspArgLeuGluAspSerAsnLysLeuLysArgSerP 400
DB      1139 TCAGTGGGCTGCACTGTCTCCGACAGACAGCTGGAAAGCAGAAATMACTCAAGGTTCTT 1198

QY      400 heSerLeuAspLleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyP 420
DB      1199 TCTCTGGAATATCAAAATCAGTTTCATATTCAAGCAGACATGGACATCTTACATGGCT 1258

QY      420 heSerSerSerGluAspAlaLeuGluLysTyrLysProSerThrThrLeuAspGlyThra 440
DB      1259 TCTCTCATCAAGAAAGATCTTTGGAAATCATCAAACTTCACATCTCTGAGTGGAGCCA 1318

QY      440 snLysLeuCyGlnPheSerProValGlnLysLeuSerGluGlnThrProGluThrSerP 460
DB      1319 ACAAGTATGCGAGTTCTCCCTGTTTCAGGAACATTCAGAGCACTCCCGAAACAGATC 1378

QY      460 roAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAspSerG 480
DB      1379 CTGATTAAGAGAGAAAGCAGCATCCCAAGAACTGCAGACCGCCAGGCTTCAGACAGCC 1438

QY      480 InSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeuL 500
DB      1439 AGAGCAAGCATTTGCACTTCGCTCAGAACAGAGCAGTGCACCGCCAGAGGCTCCCTTT 1498

QY      500 euSerProLeuHisArgSerGlySerValGlnAspAsnTyrHisThrSerPheLeuPheG 520
DB      1499 TATCTCACTGATCGAAGAGTGGAGCGCTGAGAGACATTAACACACCACTTCCTTTTG 1558

QY      520 LysLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLysLeuLysGlyTyrPH 540
DB      1559 GCCCTTCCACAGCCAGCAGCACTCAGCAAGATCTGTGGCTGGGCTTTAAGGGGTGGC 1618

QY      540 isSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyrPheA 560
DB      1619 ACTCGATATCTTGGGCCCCCGAGACCTTCACTCCCTTCCCTGACAGCACTGATTTTG 1678

QY      560 LathrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysSerAlaSerTyrS 580
DB      1679 CCACAGATGCTCACACTTCTACTGCTGCTCAGCATCATCGAGAGCAGTGCAGATTACT 1738

QY      580 eAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArgArg 600
DB      1739 CTGCTTACGCTGACAGCAGCTGCGCACTTGCGGAGACCAAACTATTAATCTGTGCGCAGCG 1798

QY      600 rgGluLysProSerAspArgAlaAspSerArgArgSerTyrHisGluGluLysProPheG 620
  
```

|||||
Db 1799 GCGAAGCAAGTGAAGAGTCTGCGGGGAGCTGTCATGAAGAGCCCTTTG 1858
|||
Qy 620 TULYSGINPHELYSARGTSERCYGLIMETGLUPHEGLYGLUSERIEMETSERGLU 640
|||
Db 1859 AAAAGCATTTAAACCCAGAGCTGCCAATGGAAATTTGGAGAGACATCATGTACAGCA 1918
|||
Qy 640 GAARGSERARGTUGLULYGLYLVYVALGLYSEGLINSERSEPHESERGLYSEMETG 660
|||
Db 1919 ACAGGTACAGGAGAGCTGGGGAAGTGGGAGTCACTTTCGGGAGAGATGG 1978
|||
Qy 660 TULLEGLULYLSER 665
|||
Db 1979 AATCATTTGAGCTCTCC 1995
|||
RESULT 2
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589)...(2583)
US-09-816-494-1
Alignment Scores:
Pred. No.: 0 Length: 3544
Score: 472.00 Matches: 663
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 70.98% Indels: 4
DB: 4 Gaps: 0
US-10-029-345A-109 (1-665) x US-09-816-494-1 (1-3544)
Qy 1 MetAlaHhIGlUmeCtIegLYThrGlnIleValThrGluArgLeuValAlaLeuLeuGlu 20
Db 589 ATGGCCCATGAGATGATTTGAGACTCAAAATTTGACTGAGAGTGGCTCTGCTGGAA 648
|||
Qy 21 SerGluThrGluValLeuLeuIleAAsPSeRArgProPheValGluTyrAenThrSer 40
Db 649 AGTGAAGCGAAGAAAGTGTCTTAATGATAGCCGGCCATTGGGATTCATAATACATCC 708
|||
Qy 41 HisIleLeuGluAlaIleAenIleAAsPSeRArgLeuLeuMetLysArgArgLeuGlnGln 60
Db 709 CACATTTTGGAAACCATTAATATCACTGCTCCAGCTTATGAAGCAAGAGTTGCCAAG 768
|||
Qy 61 AspLysValLeuIleThrGluLeuIleGlnHisSerAlaLysHisLysValAAsPLeAsP 80
Db 769 GACAAAGTGTAAATTCAGAGCTCATCCAGCATTCAGCGAACAATAGGTTGACATTTGAT 828
|||
Qy 81 CysSerGlnLysValValTyrAAsPGLNSeRSeRGLNAsPValAlaSerLeuSerSer 100
Db 829 TGCAGTCAGAAAGTTGATTTACGATCAGAGCTCCCAAGATGTTGCTCTCTCTTCA 888
|||
Qy 101 AspCysPheLeuThrValLeuLeuGluLysLeuGluLysSerPheAAsPValHisLeu 120
Db 889 GACGTTTCTCACTGTAATCTTGGGATTAATCGAAGAAAGAGCTTCAACTCTGTTCACTG 948
|||

Qy 121 LeuAlaGlyLysPheAlaGluPheSerArgCysPheProGlyLeuGlyGluLysSer 140
Db 949 CTTCAGAGTGGGTTTCTGATGTTCTCTGTTGTTCCCTGCTCTGTAAGAAATATCC 1008
|||
Qy 141 ThrLeuValProThrCysAlleSerGlnProCysLeuProValAlaAenIleGlyProThr 160
Db 1009 ACTCTAGTCCCTACCTGACATTTCTCAGCTTCTTCACTGTTGCCAACAATTTGGCCAAAC 1068
|||
Qy 161 ArgIleLeuProAAsPLeuThrLeuGlyCysGlnArgAspValLeuAAsPLeuLeu 180
Db 1069 CGAATTTCTCCAAATCTTATCTTGGCTCCAGCGAGATGCTCTCAACAAGAGCTGAT- 1127
|||
Qy 181 -GlnGlnAAsGlyIleGlyTyrValLeuAAsPLeuSerTyr-ThrCysProLysProAsp 200
Db 1128 GACAGCAAGATGGATTTGTTATGTGTTAAATGCCAGCA- TACCTGTCCAAAGCTGACT 1186
|||
Qy 200 HeIleProGluSerHisPheLeuArgValProValAAsAAsPSeRPhCyGluLysIle 220
Db 1187 TTATCCCGAGTCTGATTTCTGCGGTGCTGCTGTAATGACAGCTTTGTGAGAAATTT 1246
|||
Qy 220 euProTTrPLeuAAsPLeuSerValAAsPheIleGluLysAlaLysAAsPLeuGlyCys 240
Db 1247 TGCCTGCTTGGACAAATCAGATGATTTCAATTGAGAAACCAAGCCTCAATGGATGTG 1306
|||
Qy 240 AlLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIle 260
Db 1307 TTTGATGTCAGCTTTAGCTGGAGATCCCGCTCCGACCACTGCTATATGCTTACATCA 1366
|||
Qy 260 eLysArgMetAAsPSeRLeuAAsPLeuAlaTyrArgPheValLysGluLysArgPro 280
Db 1367 TGAAGAGATGGACATGCTTTAGATGAGAGCTTACAGATTTGGAAAGAAAGAGCCTTA 1426
|||
Qy 280 hrIleSerProAAsPheAAsPheLeuGlyGlnLeuLeuAAsPTrGluLysIleLysA 300
Db 1427 CTATATCTCCAAACTCAATTTTCTGGGCCCACTCTCGACATATGAGAAAGATTAAG 1486
|||
Qy 300 enGlnThrGlyLaseRGLYProLysSerLysLeuLysLeuLysHisLeuGluLysPro 320
Db 1487 ACCAGAGTGAAGATAGAGGCCAAGAGCAAACTCAAGCTGTGCACTGGAGAACCA 1546
|||
Qy 320 enGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeuSerPro 340
Db 1547 ATGAACCTTCTCCTGCTGCTGCTCAGAGGTGGAGCAGAAAGCGAGAGCCCTCAGTCCAC 1606
|||
Qy 340 roCysAlaAsPSeRAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSer 360
Db 1607 CCGTGTCCGACTGCTGCTACCTCAGAGGACAGCAAGAGCCGCTGCAATCCGCCAGCG 1666
|||
Qy 360 AlProSerValProSerValGlnProSerLeuLeuGluAAsPSeRProLeuValGlnAla 380
Db 1667 TGCCACGCGTCCAGCGTGCAGCCGCTGCTGTTAGAGACAGCCGCTGTGTAACAGCGC 1726
|||
Qy 380 euSerGlyLeuHisLeuSerAlaAAsPArgLeuGluAAsPSeRAsnLysLeuLysArgSer 400
Db 1727 TCAGTGGCTGCACCTGCTCCGACAGACAGCTGGAAACAGCAATAGCTCAACGTTCT 1786
|||
Qy 400 heSerLeuAAsPLeuLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGly 420
Db 1787 TCTCTGTGATATCAATACAGTTTCATATTACGCCAGTGGCAGCATCTTCAATGGCT 1846
|||
Qy 420 heSerSerSerGluAAsPAlaLeuGluTyrTyrLysProSerThrThrLeuAAsPThr 440
Db 1847 TCTCTCATCAGAAAGTGTGGAAATCTTCAAAACCTTCACTGATCTGGATGGAGCA 1906
|||
Qy 440 snLysLeuGlyGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThrSer 460
Db 1907 ACAAGCTAGCCAGTCTCCTCCCTTCAAGAACTATCGAGAGCATCCCGAAGAACAGT 1966
|||
Qy 460 roAAsPArgGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSerAAsPSe 480
Db 1967 CTGATTAAGAGAGAGCCACATCCCAAGAGCTGCAGACCGCCAGAGCTTCAAGACAGCC 2026
|||
Qy 480 InSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArgSerLeu 500
|||

DB 2027 AGAGCAGAGGATTGATTGGTTCAGAACGACGAGTGGACCGCCAGAGGTCCCTT 2086
QY 500 eUSErProLeuHISaTgSerGlySerValGluApsAntYrHISrSerPheLeuPhg 520
DB 2087 TATCTCCACTGCATCCAAAGTGGAGGCTGGAGACATTAACACACACTTCTCTTGG 2146
QY 520 lYLeuSerHrSerGlnGlnHISleuThYlSerAlaGlyLeuGlyLeuYrGlyTriH 540
DB 2147 GCTTTCCACCGACCGACGACCTTACAGAGTCTCTGCTGGGCTTTAAGGCTGGC 2206
QY 540 iSSErAspIleLeuAlaProGlnHrSerThrProSerLeuHrSerSerTriYrPhea 560
DB 2207 ACTGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGACAGCTGGTATTGG 2266
QY 560 lATHrGluSerHrHISpHeYrYrSerAlaSerAlaIleYrGlyGlySerAlaSerTyrS 580
DB 2267 CCACAGAGTCTCACCTTCTACTCTGCTCAGCCATCTACGAGGAGCACTGCGCACTTACT 2326
QY 580 eRAlATySerCySerSerGlnLeuProThrCySeGlyApsGlnValYrSerValArgArg 600
DB 2327 CTGCTTACGCTGACCGACGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCGCAGGC 2386
QY 600 rGGLnYrProSerAspArgAlaAspSerArgYrSerTriPheGlyGlnYrSerProPhg 620
DB 2387 GCGAGAGGCAAGTGCAGAGCTGACTCGCGGCGAGCTGGCATGAGAGAGCCCTTGG 2446
QY 620 lULySGlnHelysATgATgSerCySGlnMetGlnPheGlyGlnSerIleMetSerGlnA 640
DB 2447 AAAAGCAGTTTAAACGACAGAGCTGCAATGGAATTTGAGAGACATCATGTCAAGAA 2506
QY 640 sNArSerArgGlnGluLeuGlyLeuValGlySerGlnSerPheSerGlySerMetG 660
DB 2507 AAGGTTCACGGAGAGCTGGGAAAGTGGCAGTCACTTACTTTTGGGCAAGCATGG 2566
QY 660 lULleIleGlnValSer 665
DB 2567 AATCATGTGAGGTCTCC 2583
RESULT 3
US-09-920-668-3
; Sequence 3, Application US/0920668
; Patent No. 6482644
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowsett
; APPLICANT: Brett P. Monia
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 8 EXPRESSION
; FILE REFERENCE: RTS-0246
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135) .. (2012)
US-09-920-668-3
Alignment Scores:
Pred. No.: 1,55e-10 Length: 2377
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.01% Indels: 0
DB: 4 Gaps: 0
US-10-029-345A-109 (1-665) x US-09-920-668-3 (1-2377)
QY 242 VALHISCYsLeuAlaGlylIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetYls 261
DB 864 GTCCACTGTCTGGCTGGCATCTCCGCTCTGCCACCATGCATGCCTACATCATGAAG 923

RESULT 4
US-09-016-434-1135
; Sequence 1135, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2109 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1418933
US-09-016-434-1135
Alignment Scores:
Pred. No.: 0.0245 Length: 2109
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 4 Gaps: 0
US-10-029-345A-109 (1-665) x US-09-016-434-1135 (1-2109)
QY 240 VALleuValHISCYsLeuAlaGlylIleSerArgSer 251
DB 1216 GCTTGATCATGTGCTTGGCTGGCATTTAGCGCTCA 1251
RESULT 5
US-09-023-655-946
; Sequence 946, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION

```
/ NUMBER OF SEQUENCES: 1508
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/023,655
/ FILING DATE: HEREWITH
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0001 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 946:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2109 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: G1418933
/ US-09-023-655-946

Alignment Scores:
Pred. No.: 0.0245 Length: 2109
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-023-655-946 (1-2109)

QY 240 ValLeuValHicYseRgSerAlaThrIleSerArgSer 251
Db 1216 GTCTGTGATACATTGCTTGCTGGCATTAGCCGCTCA 1251

RESULT 6
US-09-702-705-801
/ Sequence 801 Application US/09702705
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tonglong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C14
/ CURRENT APPLICATION NUMBER: US/09/702,705
/ CURRENT FILING DATE: 2000-10-30
/ NUMBER OF SEQ ID NOS: 1833
```

```
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 801
/ LENGTH: 1619
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ US-09-702-705-801

Alignment Scores:
Pred. No.: 23.3 Length: 1619
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-702-705-801 (1-1619)

QY 246 AlaGlyIleSerArgSerAlaThrIle 254
Db 569 GCGGCGATCTCGCGGTGCGGCACACATC 595

RESULT 7
US-09-736-457-801
/ Sequence 801 Application US/09736457
/ Patent No. 6509448
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tonglong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ APPLICANT: Wang, Liqun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C15
/ CURRENT APPLICATION NUMBER: US/09/736,457
/ CURRENT FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 1864
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 801
/ LENGTH: 1619
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ US-09-736-457-801

Alignment Scores:
Pred. No.: 23.3 Length: 1619
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-736-457-801 (1-1619)

QY 246 AlaGlyIleSerArgSerAlaThrIle 254
Db 569 GCGGCGATCTCGCGGTGCGGCACACATC 595

RESULT 8
US-09-614-124B-801
/ Sequence 801 Application US/09614124B
/ Patent No. 6630574
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tonglong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
```

```
/ APPLICANT: Carter, Darriek
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ FILE REFERENCE: 210121.478C9
/ CURRENT APPLICATION NUMBER: US/09/614.124B
/ CURRENT FILING DATE: 2001-07-11
/ NUMBER OF SEQ ID NOS: 1668
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 801
/ LENGTH: 1619
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-09-614-124B-801

Alignment Scores:
Pred. No.: 23.3 Length: 1619
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-614-124B-801 (1-1619)

QY 246 AlaGlyIleSerArgSerAlaThrIle 254
DB 569 GCGGCATCTCGCGGTGCGCCACCATC 595

RESULT 9
US-09-671-325-801
/ Sequence 801, Application US/09671325
/ Patent No. 6667154
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darriek
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C12
/ CURRENT APPLICATION NUMBER: US/09/671.325
/ CURRENT FILING DATE: 2000-09-26
/ NUMBER OF SEQ ID NOS: 1825
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 801
/ LENGTH: 1619
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-09-671-325-801

Alignment Scores:
Pred. No.: 23.3 Length: 1619
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-671-325-801 (1-1619)

QY 246 AlaGlyIleSerArgSerAlaThrIle 254
DB 569 GCGGCATCTCGCGGTGCGCCACCATC 595

RESULT 10
US-09-589-184-801

/ Sequence 801, Application US/09589184
/ Patent No. 666447
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darriek
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ FILE REFERENCE: 210121.478C8
/ CURRENT APPLICATION NUMBER: US/09/589.184
/ CURRENT FILING DATE: 2000-06-05
/ NUMBER OF SEQ ID NOS: 827
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 801
/ LENGTH: 1619
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-09-589-184-801

Alignment Scores:
Pred. No.: 23.3 Length: 1619
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-589-184-801 (1-1619)

QY 246 AlaGlyIleSerArgSerAlaThrIle 254
DB 569 GCGGCATCTCGCGGTGCGCCACCATC 595

RESULT 11
US-08-990-379-1
/ Sequence 1, Application US/08990379
/ Patent No. 5998188
/ GENERAL INFORMATION:
/ APPLICANT: Stock, Philip J
/ APPLICANT: Mitera-Press, Anita
/ TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
/ FILE REFERENCE: 4104-000322USA
/ CURRENT APPLICATION NUMBER: US/08/990.379
/ CURRENT FILING DATE: 1997-12-15
/ EARLIER APPLICATION NUMBER: PCT/US96/10402
/ EARLIER FILING DATE: 1996-06-14
/ EARLIER APPLICATION NUMBER: 60/000.263
/ EARLIER FILING DATE: 1995-06-16
/ NUMBER OF SEQ ID NOS: 19
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 1987
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
US-08-990-379-1

Alignment Scores:
Pred. No.: 28.6 Length: 1987
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-109 (1-665) x US-08-990-379-1 (1-1987)

QY 246 AlaGlyIleSerArgSerAlaThrIle 254
```


Db 1218 GCCGCATCTCTAGATCAGCACCATC 1244
RESULT 12
US-08-990-379-2
Sequence 2, Application US/08990379
Patent No. 5998188
GENERAL INFORMATION:
APPLICANT: Stock, Philip J
TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
FILE REFERENCE: 4104-000322USA
CURRENT APPLICATION NUMBER: US/08/990,379
CURRENT FILING DATE: 1997-12-15
EARLIER APPLICATION NUMBER: PCT/US96/10402
EARLIER FILING DATE: 1996-06-14
EARLIER APPLICATION NUMBER: 60/000,263
EARLIER FILING DATE: 1995-06-16
NUMBER OF SEQ ID NOS: 19
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 2
LENGTH: 1993
TYPE: DNA
ORGANISM: Rattus norvegicus
US-08-990-379-2
Alignment Scores:
Pred. No.: 28.7 Length: 1993
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 2 Gaps: 0
US-10-029-345A-109 (1-665) x US-08-990-379-2 (1-1993)
QY 246 AlaGlyIleSerArgSerAlaThrIle 254
Db 1224 GCCGCATCTCTAGATCAGCACCATC 1250
RESULT 13
US-09-016-434-1291
Sequence 1291, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/016,434
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1291:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 929980
US-09-016-434-1291
Alignment Scores:
Pred. No.: 28.8 Length: 2000
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0
US-10-029-345A-109 (1-665) x US-09-016-434-1291 (1-2000)
QY 246 AlaGlyIleSerArgSerAlaThrIle 254
Db 1011 GCAGGCATTTCCGCTCAGCACCATC 1037
RESULT 14
US-09-702-705-825
Sequence 825, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darlick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 825
LENGTH: 2064
TYPE: DNA
ORGANISM: Homo sapiens
US-09-702-705-825
Alignment Scores:
Pred. No.: 29.7 Length: 2064
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0
US-10-029-345A-109 (1-665) x US-09-702-705-825 (1-2064)
QY 246 AlaGlyIleSerArgSerAlaThrIle 254
Db 1240 GCCGCATCTCTAGATCAGCACCATC 1266
RESULT 15
US-09-736-457-825

```
/ Sequence 825, Application US/09736457
/ Patent No. 6509448
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ APPLICANT: Wang, Aijun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C15
/ CURRENT APPLICATION NUMBER: US/09/736,457
/ CURRENT FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 1864
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 825
/ LENGTH: 2064
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-736-457-825

Alignment Scores:
Pred. No.: 29.7 Length: 2064
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-736-457-825 (1-2064)
Oy 246 AlaglylleserArgeraIathrille 254
Db 1240 GCGGGCATCTCGCGTGGCCACCATC 1266

RESULT 16
US-09-614-124B-825
/ Sequence 825, Application US/09614124B
/ Patent No. 6630574
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ FILE REFERENCE: 210121.478C9
/ CURRENT APPLICATION NUMBER: US/09/614,124B
/ CURRENT FILING DATE: 2001-07-11
/ NUMBER OF SEQ ID NOS: 1668
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 825
/ LENGTH: 2064
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-614-124B-825

Alignment Scores:
Pred. No.: 29.7 Length: 2064
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0
```

```
US-10-029-345A-109 (1-665) x US-09-614-124B-825 (1-2064)
Oy 246 AlaglylleserArgeraIathrille 254
Db 1240 GCGGGCATCTCGCGTGGCCACCATC 1266

RESULT 17
US-09-671-325-825
/ Sequence 825, Application US/09671325
/ Patent No. 6667154
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C12
/ CURRENT APPLICATION NUMBER: US/09/671,325
/ CURRENT FILING DATE: 2000-09-26
/ NUMBER OF SEQ ID NOS: 1825
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 825
/ LENGTH: 2064
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-671-325-825

Alignment Scores:
Pred. No.: 29.7 Length: 2064
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-671-325-825 (1-2064)
Oy 246 AlaglylleserArgeraIathrille 254
Db 1240 GCGGGCATCTCGCGTGGCCACCATC 1266

RESULT 18
US-09-589-184-825
/ Sequence 825, Application US/09589184
/ Patent No. 6686447
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ FILE REFERENCE: 210121.478C8
/ CURRENT APPLICATION NUMBER: US/09/589,184
/ CURRENT FILING DATE: 2000-06-05
/ NUMBER OF SEQ ID NOS: 827
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 825
/ LENGTH: 2064
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-589-184-825
```

Alignment Scores:

Pred. No.:	29.7	Length:	2064
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.35%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345A-109 (1-665) x US-09-589-184-825 (1-2064)

QY 246 AAGlyIleSerArgSerAlaThrIle 254
DB 1240 GCGGCATCTCGCGGTCCGCCACCATC 1266

RESULT 19

US-09-702-705-826
Sequence 826, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702.705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 826
LENGTH: 2109
TYPE: DNA
ORGANISM: Homo sapiens
US-09-702-705-826

Alignment Scores:

Pred. No.:	30.4	Length:	2109
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.35%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345A-109 (1-665) x US-09-702-705-826 (1-2109)

QY 246 AAGlyIleSerArgSerAlaThrIle 254
DB 1200 GCGGCATCTCGCGGTCCGCCACCATC 1226

RESULT 20

US-09-736-457-826
Sequence 826, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 826
LENGTH: 2109
TYPE: DNA
ORGANISM: Homo sapiens
US-09-736-457-826

Alignment Scores:

Pred. No.:	30.4	Length:	2109
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.35%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345A-109 (1-665) x US-09-736-457-826 (1-2109)

QY 246 AAGlyIleSerArgSerAlaThrIle 254
DB 1200 GCGGCATCTCGCGGTCCGCCACCATC 1226

RESULT 21

US-09-614-124B-826
Sequence 826, Application US/09614124B
Patent No. 6630574
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.478C9
CURRENT APPLICATION NUMBER: US/09/614,124B
CURRENT FILING DATE: 2001-07-11
NUMBER OF SEQ ID NOS: 1668
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 826
LENGTH: 2109
TYPE: DNA
ORGANISM: Homo sapiens
US-09-614-124B-826

Alignment Scores:

Pred. No.:	30.4	Length:	2109
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.35%	Indels:	0
DB:	4	Gaps:	0

US-10-029-345A-109 (1-665) x US-09-614-124B-826 (1-2109)

QY 246 AAGlyIleSerArgSerAlaThrIle 254
DB 1200 GCGGCATCTCGCGGTCCGCCACCATC 1226

RESULT 22

US-09-671-325-826
Sequence 826, Application US/09671325
Patent No. 6667154
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.

```

/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Pan, Liqun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.478C12
/ CURRENT APPLICATION NUMBER: US/09/671.325
/ NUMBER OF SEQ ID NOS: 1825
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 826
/ LENGTH: 2109
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-671-325-826

```

```

Alignment Scores:
Pred. No.: 30.4 Length: 2109
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

```

US-10-029-345A-109 (1-665) x US-09-671-325-826 (1-2109)

OY 246 AlaglytleSerArgeSerAlaThrIle 254

Db 1200 GCGGGCATCTCGCGGTGCGCCACCATC 1226

```

RESULT 23
US-09-589-184-826
/ Sequence 826, Application US/09589184
/ Patent No. 6686447
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tonglong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ FILE REFERENCE: 210121.478C8
/ CURRENT APPLICATION NUMBER: US/09/589,184
/ NUMBER OF SEQ ID NOS: 827
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 826
/ LENGTH: 2109
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-589-184-826

```

```

Alignment Scores:
Pred. No.: 30.4 Length: 2109
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

```

US-10-029-345A-109 (1-665) x US-09-589-184-826 (1-2109)

OY 246 AlaglytleSerArgeSerAlaThrIle 254

Db 1200 GCGGGCATCTCGCGGTGCGCCACCATC 1226

```

RESULT 24
US-09-016-434-1100
/ Sequence 1100, Application US/09016434
/ Patent No. 6500938
/ GENERAL INFORMATION:
/ APPLICANT: Janice Au-Young
/ APPLICANT: Jeffrey J. Seilhamer
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
/ NUMBER OF SEQUENCES: 1490
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/016,434
/ FILING DATE: HERewith
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0002 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 1100:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2240 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: g1255784
US-09-016-434-1100

Alignment Scores:
Pred. No.: 32.3 Length: 2240
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-016-434-1100 (1-2240)

OY 246 AlaglytleSerArgeSerAlaThrIle 254
Db 963 GCGGGCATCTCGCGGTGCGCCACCATC 989

RESULT 25
US-09-702-705-804
/ Sequence 804, Application US/09702705
/ Patent No. 6504010
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tonglong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick

```

APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 804
LENGTH: 4637
TYPE: DNA
ORGANISM: Homo sapiens
US-09-702-705-804

Alignment Scores:
Pred. No.: 67.3 Length: 4637
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-702-705-804 (1-4637)

QY 246 AAGAGYIIeserArgeralathrile 254
DB 569 GCGGCATCTCGCGTCGCGCACCATC 595

RESULT 26
US-09-736-457-804
Sequence 804, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aljun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 804
LENGTH: 4637
TYPE: DNA
ORGANISM: Homo sapiens
US-09-736-457-804

Alignment Scores:
Pred. No.: 67.3 Length: 4637
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-736-457-804 (1-4637)

QY 246 AAGAGYIIeserArgeralathrile 254
DB 569 GCGGCATCTCGCGTCGCGCACCATC 595

RESULT 27

US-09-614-124B-804
Sequence 804, Application US/09614124B
Patent No. 6630574
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.478C9
CURRENT APPLICATION NUMBER: US/09/614,124B
CURRENT FILING DATE: 2001-07-11
NUMBER OF SEQ ID NOS: 1668
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 804
LENGTH: 4637
TYPE: DNA
ORGANISM: Homo sapiens
US-09-614-124B-804

Alignment Scores:
Pred. No.: 67.3 Length: 4637
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-614-124B-804 (1-4637)

QY 246 AAGAGYIIeserArgeralathrile 254
DB 569 GCGGCATCTCGCGTCGCGCACCATC 595

RESULT 28
US-09-671-325-804
Sequence 804, Application US/09671325
Patent No. 6667154
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C12
CURRENT APPLICATION NUMBER: US/09/671,325
CURRENT FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 1825
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 804
LENGTH: 4637
TYPE: DNA
ORGANISM: Homo sapiens
US-09-671-325-804

Alignment Scores:
Pred. No.: 67.3 Length: 4637
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

```
US-10-029-345A-109 (1-665) x US-09-671-325-804 (1-4637)
Qy 246 AlagylleserArgserAlathrlle 254
Db 569 GCGGCATCTCGCGTCGCGCACCATC 595

RESULT 29
US-09-589-184-804
; Sequence 804, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Iodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; NUMBER OF FILING DATE: 2000-06-05
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 804
; LENGTH: 4637
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-589-184-804

Alignment Scores:
Pred. No.: 67.3 length: 4637
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.35% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-589-184-804 (1-4637)
Qy 246 AlagylleserArgserAlathrlle 254
Db 569 GCGGCATCTCGCGTCGCGCACCATC 595

RESULT 30
PCT-US92-08090-3
; Sequence 3, Application PC/TUS9208090
; GENERAL INFORMATION:
; APPLICANT: Tang, J. N.
; TITLE OF INVENTION: Fusion Protein Genes for Treatment of
; TITLE OF INVENTION: Aids
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08090
; FILING DATE: 19920922
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.

US-10-029-345A-109 (1-665) x PCT-US92-08090-3 (1-1421)
Qy 605 AsparGAlaAspSerArgArGser 612
Db 312 GATCGCGCTGACTCAAGAGAGAC 335

RESULT 31
US-08-284-465-2
; Sequence 2, Application US/08284465
; Patent No. 5830457
; GENERAL INFORMATION:
; APPLICANT: GICQUEL, BRIGITTE
; APPLICANT: TIMM, JULIANO
; APPLICANT: TRIAS, JOAQUIM
; APPLICANT: DUEZ, COLETTE
; APPLICANT: PERILLI, MARIA-GRAZIA
; APPLICANT: DUSART, JEAN
; APPLICANT: FREBE, JEAN-MARIE
; TITLE OF INVENTION: RECOMBINANT BETA-LACTAMASE, USABLE AS
; TITLE OF INVENTION: CARRIER MOLECULE FOR THE PREPARATION OF IMMUNOGENIC
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF129
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1421 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Epithelial
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..6
OTHER INFORMATION: /note= "Restriction site"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1275..1280
OTHER INFORMATION: /note= "Restriction site"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1416..1421
OTHER INFORMATION: /note= "Restriction site"
PCT-US92-08090-3

Alignment Scores:
Pred. No.: 219 length: 1421
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 5 Gaps: 0
```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,465
FILING DATE: 14-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA: FR 9201713
APPLICATION NUMBER: FR 9201713
FILING DATE: 14-FEB-1992
PRIOR APPLICATION DATA: PCT/FR 93/00151
APPLICATION NUMBER: 12-FEB-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1426 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 395..1276
US-08-284-465-2

Alignment Scores:
Pred. No.: 220 Length: 1426
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 2 Gaps: 0

US-10-029-345A-109 (1-665) x US-08-284-465-2 (1-1426)

Qy 549 SerTh-ProserLeuthrSerSer 556
Db 268 TCNACGCCGCTGCTCACGACGTTG 291

RESULT 32
US-08-284-465-7
Sequence 7, Application US/08284465
Patent No. 5830457
GENERAL INFORMATION:
APPLICANT: GICQUEL, BRIGITTE
APPLICANT: TIMM, JULIANO
APPLICANT: TRIAS, JOAQUIM
APPLICANT: DUEZ, COLETTE
APPLICANT: PERILLI, MARIA-GRAZIA
APPLICANT: DUSART, JEAN
APPLICANT: FRERE, JEAN-MARIE
TITLE OF INVENTION: RECOMBINANT BETA-LACTAMASE, USABLE AS
TITLE OF INVENTION: CARRIER MOLECULE FOR THE PREPARATION OF IMMUNOGENIC
TITLE OF INVENTION: COMPOSITIONS
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,465
FILING DATE: 14-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 9201713
FILING DATE: 14-FEB-1992
PRIOR APPLICATION DATA: PCT/FR 93/00151
APPLICATION NUMBER: 12-FEB-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1426 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 395..1276
FEATURE:
NAME/KEY: misc_feature
LOCATION: 563..568
FEATURE:
NAME/KEY: misc_feature
LOCATION: 672..677
FEATURE:
NAME/KEY: misc_feature
LOCATION: 506..532
FEATURE:
NAME/KEY: misc_feature
LOCATION: 539..562
FEATURE:
NAME/KEY: misc_feature
LOCATION: 572..589
FEATURE:
NAME/KEY: misc_feature
LOCATION: 620..652
FEATURE:
NAME/KEY: misc_feature
LOCATION: 725..748
FEATURE:
NAME/KEY: misc_feature
LOCATION: 764..790
FEATURE:
NAME/KEY: misc_feature
LOCATION: 815..835
FEATURE:
NAME/KEY: misc_feature
LOCATION: 842..874
FEATURE:
NAME/KEY: misc_feature
LOCATION: 905..925
FEATURE:
NAME/KEY: misc_feature
LOCATION: 956..988
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1016..1036
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1052..1075
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1061..1066
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1091..1117
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1127..1156
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1172..1204

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1235..1276
US-08-284-465-7

Alignment Scores:
Pred. No.: 220      Length: 1426
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.20%      Indels: 0
DB: 2      Gaps: 0

US-10-029-345A-109 (1-665) x US-08-284-465-7 (1-1426)
QY 549 SerThProSerLeuthrSerSer 556
DB 268 TCACGCGCTGCTCACAGTTCC 291

RESULT 33
PCT-US92-08090-2
; Sequence 2, Application PC/TUS9208090
; GENERAL INFORMATION:
; APPLICANT: Tang, J. N.
; TITLE OF INVENTION: Fusion Protein Genes for Treatment of
; TITLE OF INVENTION: Aids
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08090
; FILING DATE: 19920922
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRP129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1448 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Epithelial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..6
; OTHER INFORMATION: /note= "Restriction site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1275..1280
; OTHER INFORMATION: /note= "Restriction site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1444..1448
```

```
OTHER INFORMATION: /note= "Restriction site"
PCT-US92-08090-2

Alignment Scores:
Pred. No.: 224      Length: 1448
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.20%      Indels: 0
DB: 5      Gaps: 0

US-10-029-345A-109 (1-665) x PCT-US92-08090-2 (1-1448)
QY 605 AspaGAlaapSerArgArgSer 612
DB 312 GATCGCGCTGACTCAGAGAGAGC 335

RESULT 34
US-08-236-311-5
; Sequence 5, Application US/08236311
; Patent No. 556335
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/236,311
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Haasek, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 444P1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1508 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-236-311-5

Alignment Scores:
Pred. No.: 233      Length: 1508
Score: 8.00      Matches: 8
```


Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
Gaps: 0
DB: 0

US-10-029-345a-109 (1-665) x US-08-236-311-5 (1-1508)

QY 605 AaPaRgAlaAspSerArgSer 612
DB 403 GATCGCGCTGACTCAAGAAGAGC 426

RESULT 35
US-08-457-918-5
; Sequence 5, Application US/08457918
; Patent No. 6117655
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,918
; FILING DATE: 1-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/236311
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0444PLC3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8228
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1508 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-457-918-5

Alignment Scores:
Pred. No.: 233 Length: 1508
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0

DB: 3 Gaps: 0

US-10-029-345a-109 (1-665) x US-08-457-918-5 (1-1508)

QY 605 AaPaRgAlaAspSerArgSer 612
DB 403 GATCGCGCTGACTCAAGAAGAGC 426

RESULT 36
US-08-417-495-3
; Sequence 3, Application US/08417495
; Patent No. 5843728
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Redirection of Cellular Immunity by
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,495
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,866
; FILING DATE:
; APPLICATION NUMBER: US/07/847,566
; FILING DATE:
; APPLICATION NUMBER: 07/665,961
; FILING DATE: March 7, 1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/119002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1599 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-417-495-3

Alignment Scores:
Pred. No.: 247 Length: 1599
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
Gaps: 0

US-10-029-345a-109 (1-665) x US-08-417-495-3 (1-1599)

QY 605 AaPaRgAlaAspSerArgSer 612
DB 232 GATCGCGCTGACTCAAGAAGAGC 255

RESULT 37

US-08-284-391B-3
Sequence 3, Application US/08284391B
Patent No. 5851828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1599 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-284-391B-3
Alignment Scores:
Pred. No.: 247
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.20%
Matches: 8
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0
US-10-029-345A-109 (1-665) x US-08-284-391B-3 (1-1599)
QY 605 AspaAgaAaSpSerAgaAgaSer 612
DB 232 GATCGCGCTGACTCAAGAGAAGC 255
RESULT 38
US-09-218-950-3
Sequence 3, Application US/09218950
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles

APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1599 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-218-950-3
Alignment Scores:
Pred. No.: 247
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.20%
Matches: 8
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0
US-10-029-345A-109 (1-665) x US-09-218-950-3 (1-1599)
QY 605 AspaAgaAaSpSerAgaAgaSer 612
DB 232 GATCGCGCTGACTCAAGAGAAGC 255
RESULT 39
PCT-US92-01785-3
Sequence 3, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA

COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1599 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US92-01785-3

Alignment Scores:
Pred. No.: 247 Length: 1599
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 5 Gaps: 0

US-10-029-345A-109 (1-665) x PCT-US92-01785-3 (1-1599)

QY 605 AsparGAlaasparGArGArGser 612
DB 232 GATCGCGCTGACTCAAGAGAGAC 255

RESULT 40
PCT-US95-00454-3
Sequence 3, Application PC/TUS9500454
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
TITLE OF INVENTION: Bearing Cells
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1599 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US95-00454-3

Alignment Scores:
Pred. No.: 247 Length: 1599
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 5 Gaps: 0

US-10-029-345A-109 (1-665) x PCT-US95-00454-3 (1-1599)

QY 605 AsparGAlaasparGArGArGser 612
DB 232 GATCGCGCTGACTCAAGAGAGAC 255

RESULT 41
US-08-916-902A-2/c
Sequence 2, Application US/08916902A
Patent No. 5871930
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PasteSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,902A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT04
CLONE: 927955
US-08-916-902A-2

Alignment Scores:
Pred. No.: 258 Length: 1669
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
Gaps: 0
DB: 2

US-10-029-345A-109 (1-665) x US-08-916-902A-2 (1-1669)

Qy 137 GIUGLYSserThreValPro 144
DB 1349 GAAGGAAAAAGCATTGGTCCCC 1326

RESULT 42
US-09-213-389-2/c
Sequence 2, Application US/09213389
Patent No. 5977072
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT04
CLONE: 927955

US-09-213-389-2

Alignment Scores:
Pred. No.: 258 Length: 1669
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
Gaps: 0
DB: 2

US-10-029-345A-109 (1-665) x US-09-213-389-2 (1-1669)

Qy 137 GIUGLYSserThreValPro 144
DB 1349 GAAGGAAAAAGCATTGGTCCCC 1326

RESULT 43
US-08-417-495-1
Sequence 1, Application US/08417495
Patent No. 5843728
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Redirection of Cellular Immunity by
TITLE OF INVENTION: Receptor
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,495
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/203,866
FILING DATE:
APPLICATION NUMBER: US/07/847,566
FILING DATE:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1728 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-417-495-1

Alignment Scores:
Pred. No.: 267 Length: 1728
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
Gaps: 0
DB: 2

US-10-029-345A-109 (1-665) x US-08-417-495-1 (1-1728)

QY 605 AspaArgAlaAspSerArgArgSer 612

DB 232 GATCGCGCTGACTCAAGAAGAAC 255

RESULT 44

US-08-284-391B-1

; Sequence 1, Application US/08284391B

; Patent No. 5851828

; GENERAL INFORMATION:

; APPLICANT: Seed, Brian

; APPLICANT: Banapour, Babak

; APPLICANT: Romeo, Charles

; APPLICANT: Kolanus, Waldemar

; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Clark & Elbing LLP

; STREET: 176 Federal Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/284,391B

; FILING DATE: 02-AUG-1994

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/195,395

; FILING DATE: 14-FEB-1994

; APPLICATION NUMBER: 07/847,566

; FILING DATE: 06-MAR-1992

; APPLICATION NUMBER: 07/665,961

; FILING DATE: 07-MAR-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Elbing, Karen L.

; REGISTRATION NUMBER: 35,238

; REFERENCE/DOCKET NUMBER: 00786/247001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-428-0200

; TELEFAX: 617-428-7045

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1728 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-08-284-391B-1

Alignment Scores:

Pred. No.: 267

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.20%

DB: 2

US-10-029-345A-109 (1-665) x US-08-284-391B-1 (1-1728)

QY 605 AspaArgAlaAspSerArgArgSer 612

DB 232 GATCGCGCTGACTCAAGAAGAAC 255

RESULT 45

US-09-218-950-1

; Sequence 1, Application US/09218950

; Patent No. 6284240

; GENERAL INFORMATION:

; APPLICANT: Seed, Brian

; APPLICANT: Banapour, Babak

; APPLICANT: Romeo, Charles

; APPLICANT: Kolanus, Waldemar

; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Clark & Elbing LLP

; STREET: 176 Federal Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/218,950

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/284,391

; FILING DATE: 02-AUG-1994

; APPLICATION NUMBER: 08/195,395

; FILING DATE: 14-FEB-1994

; APPLICATION NUMBER: 07/847,566

; FILING DATE: 06-MAR-1992

; APPLICATION NUMBER: 07/665,961

; FILING DATE: 07-MAR-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Elbing, Karen L.

; REGISTRATION NUMBER: 35,238

; REFERENCE/DOCKET NUMBER: 00786/247001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-428-0200

; TELEFAX: 617-428-7045

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1728 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-09-218-950-1

Alignment Scores:

Pred. No.: 267

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.20%

DB: 3

US-10-029-345A-109 (1-665) x US-09-218-950-1 (1-1728)

QY 605 AspaArgAlaAspSerArgArgSer 612

DB 232 GATCGCGCTGACTCAAGAAGAAC 255

RESULT 46

PCT-US92-01785-1

; Sequence 1, Application PC/TUS9201785

; GENERAL INFORMATION:

; APPLICANT: The General Hospital Corporation

TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
TITLE OF INVENTION: Chimeras
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1728 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US92-01785-1

Alignment Scores:
Pred. No.: 267 Length: 1728
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 5 Gaps: 0

US-10-029-345A-109 (1-665) x PCT-US92-01785-1 (1-1728)

Qy 605 AsparaGalaaspSerargSer 612
Db 232 GATCGCGCTGACTCAAGAAGAGC 255

RESULT 47
PCT-US95-00454-1
Sequence 1, Application PC/TUS9500454
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
TITLE OF INVENTION: Bearing Cells
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)

SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1728 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US95-00454-1

Alignment Scores:
Pred. No.: 267 Length: 1728
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 5 Gaps: 0

US-10-029-345A-109 (1-665) x PCT-US95-00454-1 (1-1728)

Qy 605 AsparaGalaaspSerargSer 612
Db 232 GATCGCGCTGACTCAAGAAGAGC 255

RESULT 48
US-08-466-368-3
Sequence 3, Application US/08466368
GENERAL INFORMATION:
APPLICANT: Madden, Paul J.
APPLICANT: Littman, Dan R.
APPLICANT: Chess, Leonard
APPLICANT: Axel, Richard
APPLICANT: Weiss, Robin
APPLICANT: McDougal, J. S.
TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,368
FILING DATE: 06-JUN-1995

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 24577-F1-B/JPM/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1742 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 76..1449
;
US-08-466-368-3
;
Alignment Scores:
Pred. No.: 269 Length: 1742
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 3 Gaps: 0

US-10-029-345A-109 (1-665) x US-08-466-368-3 (1-1742)
QY 605 AsparGAlaasPserArgArgSer 612
DB 307 GATCGCGCTGACTCAAGAGAAGC 330

RESULT 49
US-09-517-605-7
; Sequence 7, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INFO
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: Homo sapiens
;
US-09-517-605-7
;
Alignment Scores:
Pred. No.: 269 Length: 1742
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-09-517-605-7 (1-1742)
QY 605 AsparGAlaasPserArgArgSer 612
DB 307 GATCGCGCTGACTCAAGAGAAGC 330

RESULT 50
US-08-470-998-1
; Sequence 1, Application US/08470998
```

```
; Patent No. 6570000
; GENERAL INFORMATION:
; APPLICANT: Maddon, Paul J.
; APPLICANT: Littman, Dan R.
; APPLICANT: Chase, Leonard
; APPLICANT: Axel, Richard
; APPLICANT: Weiss, Robin
; APPLICANT: McDougal, J. S.
; TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
; TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Cooper & Dunham LLP
; STREET: 1185 Avenue of Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
;
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,998
FILING DATE: 06-JUN-1995
;
CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 24577-F1-B/JPM/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1742 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 76..1449
;
US-08-470-998-1
;
Alignment Scores:
Pred. No.: 269 Length: 1742
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.20% Indels: 0
DB: 4 Gaps: 0

US-10-029-345A-109 (1-665) x US-08-470-998-1 (1-1742)
QY 605 AsparGAlaasPserArgArgSer 612
DB 307 GATCGCGCTGACTCAAGAGAAGC 330

Search completed: June 22, 2004, 10:11:37
Job time : 156 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 14:11:03 ; Search time 60 Seconds
(without alignments)
3131.566 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665

Sequence: 1 MAHEMIGTQIVTERLVALLE.....LGKVSQSSEFSGSMETIEVS 665

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 473

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database :

A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	665	100.0	665	5 ABR52381	ABR52381 Protein r
2	665	100.0	665	5 ABR52407	ABR52407 Protein r

ALIGNMENTS

RESULT 1
ABR52381
ID ABR52381 standard; protein; 665 AA.
XX
AC ABR52381;
XX
DT 19-JUN-2003 (first entry)
DE Protein relating to the invention SEQ ID NO: 109.
XX
XX antiproliferative; hepatotropic; nephrotropic; antiarthritic;
KM antiproliferative; cardiant; cytotatic; gene therapy; liver disease;
KM proliferative disorder; renal failure; cardiovascular disorder;
KM immunological disorder; arthritis; psoriasis; congenital heart defect;
KM congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
OS Homo sapiens.
XX

PN WO200257460-A2.
XX
XX 25-JUL-2002.
XX
XX 20-DEC-2001; 2001WO-US050459.
XX
XX 20-DEC-2000; 2000US-0256868P.
PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
PI Siemers N, Bol D, Schieven G, Finger J, Todderud CG, Basolino D;
PI Krystek S, Mcatee P, Suchard S, Banas D;
XX
XX WPI; 2002-599721/64.
DR N-PSDB; ACC60559.
XX
XX Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
XX Claim 5; Fig 12; 801pp; English.
XX
XX The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
CC antirheumatic, antipsoriatic, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX
XX
SQ Sequence 665 AA;
Query Match 100.0%; Score 665; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAHEMIGTQIVTERLVALLESGTEKVLIDSRPVEYNSHILEAININCSKUMRRLOQ 60
DB 1 MAHEMIGTQIVTERLVALLESGTEKVLIDSRPVEYNSHILEAININCSKUMRRLOQ 60
QY 61 DKVLITELIQHAKHKVDIDCSQKVVVYDQSSQDVASLSSDCEFLVTLGKLEKSPNSVHL 120
DB 61 DKVLITELIQHAKHKVDIDCSQKVVVYDQSSQDVASLSSDCEFLVTLGKLEKSPNSVHL 120
QY 121 LAGGFAPFRCPPGLCEGKSTLVPTCISQPCLPVANIGTRILPMLYLGCQRDVINKELI 180
DB 121 LAGGFAPFRCPPGLCEGKSTLVPTCISQPCLPVANIGTRILPMLYLGCQRDVINKELI 180
QY 181 QONGIGYVUNASVYTCRPPDFIESHFLRPVNDSPCEKLLPMLDKSVDTIEKAKSNGCV 240
DB 181 QONGIGYVUNASVYTCRPPDFIESHFLRPVNDSPCEKLLPMLDKSVDTIEKAKSNGCV 240
QY 241 LVHCLAGISRSATIAIAYIMKMDMSLDAYRFVKEKPTISPNFPLQLLDYEKKIN 300
DB 241 LVHCLAGISRSATIAIAYIMKMDMSLDAYRFVKEKPTISPNFPLQLLDYEKKIN 300
QY 301 QTGASGPKSKLKLHLKEKNEPVPVAVSEGGQKSETPLPSPCADSATSEAAQRPVHPASV 360
DB 301 QTGASGPKSKLKLHLKEKNEPVPVAVSEGGQKSETPLPSPCADSATSEAAQRPVHPASV 360
QY 361 PSVPSVQPSLLDPSFLVQALSGLHLSADRLDPSNKLKRSFLDISVSYSASMAASLHGF 420

```
DB 361 PSVPSVQPSLLEDSPLVQALSGHLISADRLBDSNKLKRSFSLDIKSVSYSASMAASLHG 420
QY 421 SSSDALLEYKPSSTLLDGTNKLCOFSPVOELSEQTPETSPDKEASIPKKLQTARPSDQ 480
DB 421 SSSDALLEYKPSSTLLDGTNKLCOFSPVOELSEQTPETSPDKEASIPKKLQTARPSDQ 480
QY 481 SKRLHSVRTSSSGTQORSLISPLHRSQGVEDNYHTSPLFGLSTSOOHLTKSAGLKGWH 540
DB 481 SKRLHSVRTSSSGTQORSLISPLHRSQGVEDNYHTSPLFGLSTSOOHLTKSAGLKGWH 540
QY 541 SDILAPOQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDQVYSVRR 600
DB 541 SDILAPOQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDQVYSVRR 600
QY 601 QKPSDRAISRHWHSPEKQPKRRSCQMEFGESIMSENRREELGKVGSSSFGSME 660
DB 601 QKPSDRAISRHWHSPEKQPKRRSCQMEFGESIMSENRREELGKVGSSSFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665

RESULT 2
ABRS2407
ID ABR52407 standard; protein; 665 AA.
XX
AC ABR52407;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein relating to the invention SEQ ID NO: 148.
XX
KW antiproliferative; hepatotropic; nephrotropic; antiarthritic;
KW antipsoriatic; cardiant; cytostatic; gene therapy; liver disease;
KW proliferative disorder; renal failure; cardiovascular disorder;
KW immunological disorder; arthritis; psoriasis; congenital heart defect;
KW congestive heart failure; cancer; human; mouse; yeast; rat; fruitfly.
XX
OS Homo sapiens.
XX
PN MO200257460-A2.
XX
PD 25-JUL-2002.
XX
PF 20-DEC-2001; 2001WO-US050459.
XX
PR 20-DEC-2000; 2000US-0256668P.
PR 30-MAR-2001; 2001US-0280186P.
PR 01-MAY-2001; 2001US-0287735P.
PR 05-JUN-2001; 2001US-0295848P.
PR 25-JUN-2001; 2001US-0300465P.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Jackson DG, Feder J, Nelson T, Mintier G, Ramanathan C, Lee L;
PI Stewers N, Bol D, Schieven G, Finger J, Todderud CG, Bassolino D;
PI Krysatek S, Mcatee P, Suchard S, Banas D;
XX
XX WPI; 2002-599721/64.
DR N-PDB; ACC60572.
PT Novel polynucleotides encoding human phosphatase polypeptides useful in
PT the prevention or treatment of e.g. proliferative and cardiovascular
PT disorders.
XX
XX Disclosure; Fig 19; 801pp; English.
XX
XX The invention relates to a novel isolated nucleic acid comprising a
CC polynucleotide having a nucleotide sequence selected from 40
CC polynucleotides fully defined in the specification. The polynucleotide of
CC the invention has antiproliferative, hepatotropic, nephrotropic,
```

```
CC antiarthritic, antipsoriatic, cardiant, and cytostatic activity. The
CC polynucleotide may have a use in gene therapy. A polynucleotide or
CC polypeptide of the invention is useful for preventing, treating or
CC ameliorating a medical condition, e.g. a proliferative disorder. They are
CC also useful for treating e.g. liver disease, renal failure, immunological
CC disorders including arthritis and psoriasis, cardiovascular disorders
CC such as congenital heart defects and congestive heart failure, and
CC cancer. A method of the invention is useful for diagnosing a pathological
CC condition or susceptibility to a condition in a subject. The present
CC sequence is used in the exemplification of the invention
XX
SQ Sequence 665 AA;
Query Match 100.0%; Score 665; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAHEMIGQIVTERIVALLSEGTQKVLIDSRPFVEYNTSHLEAININCSKLMRRRLQ 60
DB 1 MAHEMIGQIVTERIVALLSEGTQKVLIDSRPFVEYNTSHLEAININCSKLMRRRLQ 60
QY 61 DKVLITTELIQHSAAKRVNDICQKVVVYDQSSQDVASLSDGFTVLIGLKEKSTNSVHL 120
DB 61 DKVLITTELIQHSAAKRVNDICQKVVVYDQSSQDVASLSDGFTVLIGLKEKSTNSVHL 120
QY 121 IAGGFAERSRCPPGICEKSTIVPTCISQPCVPANIGPTRILPMLYIGCORQVYNKEI 180
DB 121 IAGGFAERSRCPPGICEKSTIVPTCISQPCVPANIGPTRILPMLYIGCORQVYNKEI 180
QY 181 QONGIGYVLNASTYTCPKPDFIPESHFLRPVNDSECEKILPWLIDKSVDFIEKASNGCV 240
DB 181 QONGIGYVLNASTYTCPKPDFIPESHFLRPVNDSECEKILPWLIDKSVDFIEKASNGCV 240
QY 241 LVHCLAGISRSATIAIYIMKMDMSLDBATRFVKEKPTTSPNPNFGLLDYEKTKYN 300
DB 241 LVHCLAGISRSATIAIYIMKMDMSLDBATRFVKEKPTTSPNPNFGLLDYEKTKYN 300
QY 301 QTGASGPKSKLKLHLEKNEPVPVAVSFGQKSETPPLSPPCADSATSEAAQRPVHPASV 360
DB 301 QTGASGPKSKLKLHLEKNEPVPVAVSFGQKSETPPLSPPCADSATSEAAQRPVHPASV 360
QY 361 PSVPSVQPSLLEDSPLVQALSGHLISADRLBDSNKLKRSFSLDIKSVSYSASMAASLHG 420
DB 361 PSVPSVQPSLLEDSPLVQALSGHLISADRLBDSNKLKRSFSLDIKSVSYSASMAASLHG 420
QY 421 SSSDALLEYKPSSTLLDGTNKLCOFSPVOELSEQTPETSPDKEASIPKKLQTARPSDQ 480
DB 421 SSSDALLEYKPSSTLLDGTNKLCOFSPVOELSEQTPETSPDKEASIPKKLQTARPSDQ 480
QY 481 SKRLHSVRTSSSGTQORSLISPLHRSQGVEDNYHTSPLFGLSTSOOHLTKSAGLKGWH 540
DB 481 SKRLHSVRTSSSGTQORSLISPLHRSQGVEDNYHTSPLFGLSTSOOHLTKSAGLKGWH 540
QY 541 SDILAPOQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDQVYSVRR 600
DB 541 SDILAPOQSTPSTLTSSWYFATESSHFYASAIYGSASAYSAYSCQLPTCGDQVYSVRR 600
QY 601 QKPSDRAISRHWHSPEKQPKRRSCQMEFGESIMSENRREELGKVGSSSFGSME 660
DB 601 QKPSDRAISRHWHSPEKQPKRRSCQMEFGESIMSENRREELGKVGSSSFGSME 660
QY 661 IIEVS 665
DB 661 IIEVS 665
```

Search completed: June 21, 2004, 14:17:35
Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 14:15:29 ; Search time 21 Seconds

(without alignments)
3046.065 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665
Sequence: 1 MAHEMIGTQIVTERLVALLE.....LGKVGSSQSSFGSGMERIEVS 665

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 473

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database : PIR 78:*

1: d1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description

No matches found					

Search completed: June 21, 2004, 14:19:38
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 14:11:38 ; Search time 17 Seconds
(without alignments)
2036.862 Million cell updates/sec

Title: US-10-029-345A-109
Perfect score: 665
Sequence: 1 MAHEMIGTQIVTERLVALLE.....LQKVGSQSSFSGSMETIEVS 665

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 473

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

No matches found					

Search completed: June 21, 2004, 14:18:05
Job time : 17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 14:14:54 ; Search time 46 Seconds
(without alignments)
4561.296 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665

Sequence: 1 MAHEMIGTQIVTERLVALLE.....LGRVGSQSSFGSGMERIEVS 665

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 473

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description

No matches found				

Search completed: June 21, 2004, 14:19:04
JOB time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 21, 2004, 14:17:40 ; Search time 49 Seconds
(without alignments)
3831.399 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665
Sequence: 1 MAHEMIGTQIVTERLVALL.....LGKVGSSQSSFSQSMETIEVS 665

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1163542 seqs, 282313646 residues

Word size : 473

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
No matches found					

Search completed: June 21, 2004, 14:21:16
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 21, 2004, 14:16:30 ; Search time 23 Seconds
(without alignments)
1492.664 Million cell updates/sec

Title: US-10-029-345A-109

Perfect score: 665
Sequence: 1 MAHEMIGTQIVTERLVALL.....LGRVGSQSSFSGSMRIEVS 665

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 473

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 75 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile01.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

No matches found

Search completed: June 21, 2004, 14:20:14
Job time : 23 secs

This Page Blank (uspto)